

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleotide search, using frame\_plus.p2n model.

Run on: July 11, 2006, 19:46:18 ; Search time 1356.71 Seconds  
(without alignments)  
4383.493 Million cell updates/sec

Title: US-10-718-311-15  
Perfect score: 305  
Sequence: 1 MASSVISAATRSNVTOA.....LDITSIASNGRVSCHQVWH 62

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues  
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/abs/ABSSWEB spool/US10718311/runat\_11072006\_111520\_5171/app\_query.fasta\_1  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes03p  
-USER=US10718311 @CGN 1.1 8663 @runat\_11072006\_111520\_5171 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

- 1: gb\_env.\*
- 2: gb\_pat.\*
- 3: gb\_ph.\*
- 4: gb\_pl.\*
- 5: gb\_pr.\*
- 6: gb\_ro.\*
- 7: gb\_sts.\*
- 8: gb\_sy.\*
- 9: gb\_un.\*
- 10: gb\_vi.\*
- 11: gb\_ov.\*
- 12: gb\_htg.\*
- 13: gb\_in.\*
- 14: gb\_om.\*
- 15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305	100.0	684	2	AR428736 Sequence
2	305	100.0	684	2	AR455412 Sequence
3	305	100.0	684	2	AX329368 Sequence

4	305	100.0	684	2	AX382258	Sequence
5	297	97.4	599	4	TOMRBCSB	MI3543 Tomato (L.e
6	297	97.4	742	4	TOMRBCSE	M15236 Tomato RuBP
7	297	97.4	1097	4	BT013023	BT013023 Lycopersi
8	292	95.7	2776	4	LERBSC2	X05983 Tomato rbcS
9	269	88.2	204	2	AR528363	AR528363 Sequence
10	269	88.2	499	2	AR205973	AR205973 Sequence
11	269	88.2	796	4	NSRUB1	X01722 Nicotiana s
12	266	87.2	806	4	AY220079	AY220079 Nicotiana
13	262	85.9	22574	8	CVE18556	M18556 Cloning vec
14	258	84.6	979	4	TOBRUBPA	M32419 Tobacco rib
15	258	84.6	2293	4	NPRBSC8B	X13711 Nicotiana p
16	258	84.6	2293	4	TOBRBSC8B	M36685 N.plumbagin
17	258	84.6	2362	4	NTRUBS	X02353 Tobacco gen
18	257.5	84.1	546	4	POTRBCS	J03613 Potato (S.t
19	256.5	84.1	3323	4	STRBSC1	X69759 S.tuberosum
20	256	83.9	723	4	SLARBGS	L26605 Stellaria l
21	254	83.3	177	2	I24839	I24839 Sequence 8
22	254	83.3	177	2	I40228	I40228 Sequence 6
23	254	83.3	177	2	I72655	I72655 Sequence 12
24	254	83.3	225	8	ARPRECA2	X05923 Hybrid rbc-
25	254	83.3	1054	4	LERBSC3B	X05985 Tomato rbcS
26	254	83.3	4124	4	TOMRBCO	D11112 Tomato ribu
27	253.5	83.1	778	4	TOMRBCSD	M15235 Tomato RuBP
28	253	83.0	1027	4	TOMRBCSC	M13544 Tomato (L.e
29	253	83.0	1341	4	LERBSC3C	X05986 Tomato rbcS
30	253	83.0	1520	4	LERBSC3A	X05984 Tomato rbcS
31	251	82.3	1261	4	BT012936	BT012936 Lycopersi
32	249	81.6	2242	4	PERBSC08	X03820 Petunia x h
33	248	81.3	1386	4	STRBSC2C	X69762 S.tuberosum
34	248	81.3	1629	4	STRBSC2	X69760 S.tuberosum
35	247	81.0	1314	4	PERBSC11	X03821 Petunia x h
36	246.5	80.8	1598	4	STRBSC3	X69763 S.tuberosum
37	246	80.7	1703	4	STRBSC2B	X69761 S.tuberosum
38	242.5	79.5	1032	4	TOMRBCSA	M13542 Tomato (L.e
39	242.5	79.5	1454	4	LERBSC1	X05982 Tomato rbcS
40	228.5	74.9	1337	4	TOBRUBPB	M32420 Tobacco rib
41	228	74.8	801	4	AF065615	AF065615 Capsicum
42	225	73.8	546	2	CQ805080	CQ805080 Sequence
43	225	73.8	546	4	AY098970	AY098970 Arabidops
44	225	73.8	577	4	AY051025	AY051025 Arabidops
45	225	73.8	595	4	AY064686	AY064686 Arabidops

ALIGNMENTS

RESULT 1  
AR428736  
LOCUS AR428736 684 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 16 from patent US 6642036.  
ACCESSION AR428736  
VERSION AR428736.1 GI:40188466  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 684)  
AUTHORS Flint, D., Meyer, K. and Viitanen, P.  
TITLE Sinapoylglucose:malate sinapoyltransferase form malate conjugates  
from benozic acid glucosides  
JOURNAL Patent: US 6642036-A 16 04-NOV-2003;  
E. I. Du Pont de Nemours and Company; Wilmington;

WOX;

FEATURES

source

Location/Qualifiers

1..684

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 684

Score: 305.00

Matches: 62

Percent Similarity: 100.0%

Conservative: 0

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR428736 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 2
AR455412
LOCUS AX329368
DEFINITION Sequence 7 from patent US 6683231.
ACCESSION AR455412
VERSION AR455412.1 GI:42689960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Meyer, K., Viitanen, P.V. and Van Dyk, D.E.
TITLE High level production of p-hydroxybenzoic acid in green plants
JOURNAL E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
source
1. .684
/mol_type="unknown"
/organism="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR455412 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 3
AX329368
LOCUS AX329368
DEFINITION Sequence 7 from Patent WO0194607.

```

```

ACCESSION AX329368
VERSION AX329368.1 GI:18102367
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1 Meyer, K., van Dyk, D.E. and Viitanen, P.V.
AUTHORS High level production of p-hydroxybenzoic acid in green plants
TITLE Patent: WO 0194607-A 7 13-DEC-2001;
JOURNAL E.I. Dupont De Nemours (US)
FEATURES
Location/Qualifiers
source
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic CPL"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AX329368 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 4
AX382258
LOCUS AX382258
DEFINITION Sequence 16 from Patent WO0204653.
ACCESSION AX382258
VERSION AX382258.1 GI:19577037
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1 Flint, D., Meyer, K. and Viitanen, P.V.
AUTHORS Sinapolyglucose:maltate sinapolytransferase form maltate conjugates
TITLE from benzoic acid glucosides
JOURNAL Patent: WO 0204653-A 16 17-JAN-2002;
FEATURES
Location/Qualifiers
source
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="open reading frame of the chloroplast-targeted CPL fusion protein"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62

```

Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AX382258 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 1 ATGGCTTCTCTGTCATTTCTTCCAGCAGCTGTGGCCACAGCAATGTTACACAAGCT 60  
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuIleSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 61 AGCATGGTGGACCTTCTACTGGTCTCAATCTTCAGCCACTTTCCTGTTTCAAGAAG 120  
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 121 CAAACCTTGACATCACTTCCATGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180  
 |||||

QY 61 TrpHis 62  
 |||||  
 Db 181 TGGCAT 186

RESULT 5  
 TOMRBCSB 599 bp mRNA linear PLN 27-APR-1993  
 LOCUS  
 DEFINITION Tomato (L.esculentum) ribulose-1,5-bisphosphate  
 carboxylase/oxygenase (RBCS) small subunit mRNA, clone Rbcs-2A.  
 ACCESSION M13543  
 VERSION M13543.1 GI:170497  
 KEYWORDS ribulose 1,5-bisphosphate carboxylase/oxygenase.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
 AUTHORS Pichersky, E., Bernatzky, R., Tanksley, S.D. and Cashmore, A.R.  
 TITLE Evidence for selection as a mechanism in the concerted evolution of  
 Lycopersicon esculentum (tomato) genes encoding the small subunit  
 of ribulose-1,5-bisphosphate carboxylase/oxygenase  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (11), 3880-3884 (1986)  
 PUBMED 3012537  
 COMMENT Original source text: Tomato (L.esculentum) cDNA to mRNA, clone  
 Rbcs-2A.  
 Clean copy sequence for [1] kindly provided by E.Pichersky,  
 31-OCT-1986

FEATURES  
 source Location/Qualifiers  
 1..599  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 15..557  
 /notes="ribulose-1,5-bisphosphate carboxylase/ oxygenase  
 small subunit (EC 4.1.1.39)"  
 /codon\_start=1  
 /protein\_id="AAA34189.1"  
 /db\_xref="GI:170498"  
 /translation="MASSVITSSAAVATRSNVTQASVAPFTGLKSSATPVTKKQNLID  
 ITSISNGRVSCHQVWPINMKKYETLSYLPDLSDQLLSEYVLLKNGVPCLEFE  
 TENGVYRENNKSGFYDGRYTWKMLPMFGCTDQTQVLAEOEAKKAYPAQWVRIG  
 FDNVRQVCISFIAYKPEGY"

CDS  
 1..599  
 /notes="ribulose-1,5-bisphosphate carboxylase/ oxygenase  
 small subunit (EC 4.1.1.39)"  
 /codon\_start=1  
 /protein\_id="AAA34189.1"  
 /db\_xref="GI:170498"  
 /translation="MASSVITSSAAVATRSNVTQASVAPFTGLKSSATPVTKKQNLID  
 ITSISNGRVSCHQVWPINMKKYETLSYLPDLSDQLLSEYVLLKNGVPCLEFE  
 TENGVYRENNKSGFYDGRYTWKMLPMFGCTDQTQVLAEOEAKKAYPAQWVRIG  
 FDNVRQVCISFIAYKPEGY"

ORIGIN Chromosome 3, Rbcs-2 locus.  
 Alignment Scores:  
 Pred. No.: 3,99e-28 Length: 599  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSB (1-599)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 15 ATGGCTTCTCTGTCATTTCTTCCAGCAGCTGTGGCCACAGCAATGTTACACAAGCT 74  
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuIleSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 75 AGCATGGTGGACCTTCTACTGGTCTCAATCTTCCAGCCACTTTCCTGTTTCAAGAAG 134  
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 135 CAAACCTTGACATCACTTCCATGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 194  
 |||||

QY 61 Trp 61  
 |||||  
 Db 195 TGG 197

RESULT 6  
 TOMRBCSE 742 bp mRNA linear PLN 27-APR-1993  
 LOCUS  
 DEFINITION Tomato RuBP carboxylase small subunit LESS5 mRNA.  
 ACCESSION M15236  
 VERSION M15236.1 GI:170503  
 KEYWORDS carboxylase; ribulose biphosphate carboxylase.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
 AUTHORS McKnight, T.D., Alexander, D.C., Babcock, M.S. and Simpson, R.B.  
 TITLE Nucleotide sequence and molecular evolution of two tomato genes  
 encoding the small subunit of ribulose-1,5-bisphosphate carboxylase  
 JOURNAL Gene 48 (1), 23-32 (1986)  
 PUBMED 3557127  
 COMMENT Original source text: Tomato (VF36), cDNA to mRNA, clone LESS5.  
 Draft entry and computer-readable copy of sequence in [1] kindly  
 provided by T.D.McKnight (15-APR-1987).

FEATURES  
 source Location/Qualifiers  
 1..742  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 <1..742  
 /product="RuBP mRNA"  
 28..570  
 /notes="ribulose-1,5-bisphosphate carboxylase, small  
 subunit precursor"  
 /codon\_start=1  
 /protein\_id="AAA34192.1"  
 /db\_xref="GI:170504"  
 /translation="MASSVITSSAAVATRSNVTQASVAPFTGLKSSATPVTKKQNLID  
 ITSISNGRVSCHQVWPINMKKYETLSYLPDLSDQLLSEYVLLKNGVPCLEFE  
 TSHGFYVRENNKSPGYDGRYTWKMLPMFGCTDQTQVLAEOEAKKAYPAQWVRIG  
 FDNVRQVCISFIAYKPEGY"  
 28..198  
 /sig\_peptide  
 /note="ribulose-1,5-bisphosphate carboxylase signal pept"  
 199..567  
 /mat\_peptide  
 /product="ribulose-1,5-bisphosphate carboxylase"  
 614 bp upstream of Drai site.

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,27e-28 Length: 742  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSE (1-742)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||

Db 28 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACGAGCAATGTTACACAAGCT 87  
 QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIysLys 40  
 Db 88 AGCATGGTTGCACCTTCTACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTTACAAGAAG 147  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 148 CAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 207  
 QY 61 TTP 61  
 Db 208 TGG 210

RESULT 7  
 BT013023  
 LOCUS BT013023 1097 bp mRNA linear PLN 11-MAY-2004  
 DEFINITION Lycopersicon esculentum clone 114270R, mRNA sequence.  
 ACCESSION BT013023  
 VERSION BT013023.1 GI:47104438  
 KEYWORDS FLI\_CDNA.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 1097)  
 AUTHORS Kirkness,E.F., Wang,W. and Vazeille,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712  
 Medical Center Drive, Rockville, MD 20850, USA

## FEATURES

source  
 1..1097  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 /clone="114270R"  
 /tissue\_type="Pseudomonas resistant"  
 /note="TPRC058"

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,78e-28 Length: 1097  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x BT013023 (1-1097)

QY 1 MetaLaserSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 362 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACGAGCAATGTTACACAAGCT 421  
 QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIysLys 40  
 Db 422 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTTACAAGAAG 481  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 482 CAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 541  
 QY 61 TTP 61  
 Db 542 TGG 544

## RESULT 8

LERBCS2  
 LOCUS LERBCS2 2776 bp DNA linear PLN 24-FEB-1999  
 DEFINITION Tomato rbcS2 gene for ribulose 1,5-bisphosphate  
 carboxylase/oxygenase small subunit.  
 ACCESSION X05983

VERSION X05983.1 GI:19331  
 KEYWORDS rbcS gene; ribulose bisphosphate carboxylase.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 2776)  
 AUTHORS Sugita,M., Manzara,T., Pichereky,E., Cashmore,A. and Grussem,W.  
 TITLE Genomic organization, sequence analysis and expression of all five  
 genes encoding the small subunit of ribulose-1,5-bisphosphate  
 carboxylase/oxygenase from tomato

Mol. Gen. Genet. 209 (2), 247-256 (1987)  
 3478552

## JOURNAL PUBMED

REFERENCE 2 (bases 1 to 2776)

## AUTHORS

Manzara,T.

## TITLE

Direct Submission

## JOURNAL

Submitted (02-FEB-1989)

## REFERENCE

3 (bases 1 to 2776)

## AUTHORS

Manzara,T.

## TITLE

Direct Submission

## JOURNAL

Submitted (03-AUG-1989)

## COMMENT

see X05986.

## FEATURES

Location/Qualifiers

1..2776

/organism="Lycopersicon esculentum"

/mol\_type="genomic DNA"

/cultivar="VENT Cherry LA1221"

/db\_xref="taxon:4081"

/clone="Rbcs-2"

enhancer 255..265

CAAT\_signal 274..278

misc\_feature 338..349

TATA\_signal 355..365

exon 393..597

gene 421..2495

CDS join(421..597,1597..1731,1830..1882,2318..2495)

/gene="rbcS2"

/codon\_start=1

/product="ribulose 1,5-bisphosphate carboxylase/oxygenase"

/protein\_id="CAA29401.2"

/db\_xref="GI:4456641"

/db\_xref="GOA:P07179"

/db\_xref="UniProtKB/Swiss-Prot:P07179"

/translation="MASSVSSAAVATRSNVTAQSMVAPFTGLKSSATFPVTKQNLD

ITSIASNGRVSQMVPPINMKKIETLSYLPDLSDBQLSEIETLKNQWVPCLEPE

TEHGFFVRENKSGFYDGRYTWKLMFEGCTDATQVLAEVQAKKAYPQAWRIIG

FDNVRVQCISFIAYKPEGY"

421..591

/gene="rbcS2"

join(592..597,1597..1731,1830..1882,2318..2492)

/gene="rbcS2"

/product="ribulose 1,5-bisphosphate carboxylase/oxygenase"

598..1596

/number=1

/gene="rbcS2"

1597..1731

/number=2

/gene="rbcS2"

1732..1829

/number=2

/gene="rbcS2"

1739..1741

/gene="rbcS2"

/note="tca was taa in [2]"

/citation=[2]

1830..1882

/gene="rbcS2"

1883..2317

/number=3

introns



```

/ gene="rbcs2"
/ number=3
1956..1957
/ gene="rbcs2"
/ notes="aa was ada in [2]"
/ citation=[2]
2318..>2493
/ gene="rbcs2"
/ number=4
2637..2642

ORIGIN
Alignment Scores:
Pred. No.: 1.27e-26 Length: 2776
Score: 292.00 Matches: 61
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 1
Query Match: 95.7% Indels: 0
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x LERBCS2 (1-2776)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 421 ATGGCTTCCTCTGCTCACTTCTTCAGCAGCTGTTGCCACACGCAGCAATGTTACACAAGCT 480

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 481 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 540

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 541 CAAAACCTTGACATCACTTCCATTTAGTCAATGGTGGGAAGAGTTAGCTGCATGCAGGTA 600

Qy 61 TrpHis 62
Db 601 ACTCAC 606

RESULT 9
AR528363
LOCUS AR528363 204 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6723895.
ACCESSION AR528363
VERSION AR528363.1 GI:53916426
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 204)
AUTHORS DeBonte,L.R. and Shorrosh,B.S.
TITLE Plants containing a cytosolic acetyl CoA-carboxylase nucleic acid
JOURNAL Patent: US 6723895-A 3 20-APR-2004;
Cargill, Incorporated; Wayzata, MN
FEATURES
source
1..204
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.56e-25 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR528363 (1-204)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCACTTCTTCCTCTGACGAGTGGCCACCGCAGCAATGTTGCTCAAGCT 60

```

```

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTTCACCTGCGCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCGCAGCAACGGCGGAAGAGTGCATGCAATGCATGCAGGTG 180

Qy 61 Trp 61
Db 181 TGG 183

RESULT 10
AR205973
LOCUS AR205973 499 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 41 from patent US 6369296.
ACCESSION AR205973
VERSION AR205973.1 GI:21503691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 499)
AUTHORS Ratcliff,F.Giles., Martin-Hernandez,A.Montserrat. and
Raulcombe,D.Charles.
TITLE Recombinant plant viral vectors
JOURNAL Patent: US 6369296-A 41 09-APR-2002;
FEATURES
Location/Qualifiers
1..499
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.14e-24 Length: 499
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR205973 (1-499)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCCTCACTTCTTCCTCTGACGAGTGGCCACCGCAGCAATGTTGCTCAAGCT 67

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 AACATGGTTGCACCTTTCACCTGCGCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 127

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAAACCTTGACATCACTTCCATTGCGCAGCAACGGCGGAAGAGTGCATGCAATGCATGCAGGTG 187

Qy 61 Trp 61
Db 188 TGG 190

RESULT 11
NSRUB1
LOCUS NSRUB1 796 bp mRNA linear PLN 18-APR-2005
DEFINITION Nicotiana silvestris mRNA for the small subunit of ribulose
bisphosphate carboxylase.
ACCESSION X01722
VERSION X01722.1 GI:19759
KEYWORDS
SOURCE Nicotiana sylvestris
ORGANISM Nicotiana sylvestris (wood tobacco)
REFERENCE 1 (bases 1 to 796)

```

AUTHORS Pinck, M., Guille, E., Durr, A., Hoff, M., Pinck, L. and Fleck, J.  
 TITLE Complete sequence of one of the mRNAs coding for the small subunit  
 of ribulose biphosphate carboxylase of Nicotiana sylvestris  
 JOURNAL Biochimie 66 (7-8), 539-545 (1984)  
 PUBMED 6549380  
 COMMENT Data kindly reviewed (22-JUL-1985) by J. Fleck.

## FEATURES

source  
 1..796  
 /organism="Nicotiana sylvestris"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4096"  
 misc\_feature  
 1..60  
 /note="5' untranslated region"  
 repeat\_region  
 19..25  
 /note="direct repeat 1"  
 misc\_feature  
 21..46  
 /note="pot. stem and loop structure"  
 repeat\_region  
 27..33  
 /note="direct repeat 1"  
 repeat\_region  
 53..61  
 /note="direct repeat 2"  
 61..603  
 /note="unnamed protein product; small subunit protein precursor"  
 /codon\_start=1  
 /protein\_id="CAA25862.1"  
 /db\_xref="GI:19760"  
 /db\_xref="GOA:P69250"  
 /db\_xref="InterPro:IPR000894"  
 /db\_xref="UniProtKB/Swiss-Prot:P69250"  
 /translation="MASSVLSAAVATRSNVAQNMVAPFTGLKSAASFVPSRKQNL  
 ITSIAASNGRQVQCMQVPPYKPKKYYETLSYLPDLSEQLLSEVYLKNGWVPCLEFE  
 TEGFVYRENNKSPGYDGRYWTWKLPFMFGCTDATQVLAEEVEAKKAYPQAWIRIIG  
 FDNVRQVCISFIAYKPEGY"  
 transit\_peptide  
 61..231  
 /note="transit peptide (aa -57 to -1)"  
 mat\_peptide  
 232..600  
 /product="mature small subunit protein (aa 1-123)"  
 repeat\_region  
 249..257  
 /note="direct repeat 2"  
 misc\_feature  
 601..796  
 /note="3' untranslated region"  
 polyA\_site  
 796  
 /note="polyA site"

## ORIGIN

Alignment Scores:  
 Pred. No.: 21e-24 Length: 796  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x NSRUB1 (1-796)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 61 ATGGCTTCTCAGTTCTTCTCTCGACGAGTTGCGCCGAGCAATGTTGCTCAAGCT 120  
 Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40  
 Db 121 AACATGGTTGACCTTTCACCTGCGCTTAAGTCAGTCGCTCATTCCTGTTTCAAGGAA 180  
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 181 CAAACACCTTGACATCACTTCCATTGCGACCAACGCGGAGAGTGCATGTCAGGTG 240

Qy 61 Trp 61  
 Db 241 TGG 243

RESULT 12

AY220079

LOCUS

DEFINITION

Nicotiana tabacum putative ribulose biphosphate carboxylase small subunit protein precursor, mRNA, complete cds; nuclear gene for chloroplast product.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Nicotiana tabacum (common tobacco)  
 Nicotiana tabacum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (15-JAN-2003) Biochemistry and Molecular Biology, Oklahoma State University, 246 Noble Research Center, Stillwater, OK 74078, USA

FEATURES

source

1..806

/organism="Nicotiana tabacum"

/mol\_type="mRNA"

/db\_xref="taxon:4097"

61..603

/note="similar to Nicotiana silvestris ribulose biphosphate carboxylase small subunit"

/codon\_start=1

/product="putative ribulose biphosphate carboxylase small subunit protein precursor"

/protein\_id="AAP03874.1"

/db\_xref="GI:30013663"

/translation="MASSVLSAAVATRTNVAQNMVAPFTGLKSAASFVPSRKQNL  
 ITSIAASNGRQVQCMQVPPYKPKKYYETLSYLPDLSEQLLSEIYLLKNGWVPCLEFE  
 TERGFVYRENNKSPGYDGRYWTWKLPFMFGCTDATQVLAEEVEAKKAYPEAWIRIIG  
 FDNVRQVCISFIAYKPEGY"

ORIGIN

Alignment Scores:

Pred. No.: 5.14e-24 Length: 806

Score: 266.00 Matches: 52

Percent Similarity: 96.7% Conservative: 7

Best Local Similarity: 85.2% Mismatches: 2

Query Match: 87.2% Indels: 0

DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x AY220079 (1-806)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 61 ATGGCTTCTCAGTTCTTCTCTCGACGAGTTGCGCCACCAATGTTGCTCAAGCT 120

Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40

Db 121 AACATGGTTGACCTTTCACCTGCGCTTAAGTCAGTCGCTCATTCCTGTTTCAAGGAA 180

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 181 CAAACACCTTGACATCACTTCCATTGCTAGCAATGTTGAAGAGTGCATGTCAGGTA 240

Qy 61 Trp 61

Db 241 TGG 243

RESULT 13

CVE18556

LOCUS

DEFINITION

Cloning vector pSLJ8313, T-DNA region.

ACCESSION

Y18556

22574 bp DNA

linear

SYN 24-FEB-1999



```

/misc_feature
/note="source: Cauliflower mosaic virus"
14508..14595
/note="5' leader sequence, Tobacco mosaic virus"
14596..20106
/note="En/Spm transposable element, defective"
15790..17439
/note="artificial bar gene"
15790..16698
/gene="bar, artificial"
15790..16074
/gene="bar, artificial"
/note="source: nopaline synthase gene, Agrobacterium tumefaciens, plasmid pTi37"
16147..16698
/gene="bar, artificial"
/note="source: Streptomyces hygroscopicus"
/codon_start=1
/transl_table=11
/product="phosphothic acetyl transferase"
/protein_id="CAA77216.1"
/db_xref="GI:4191251"
/translation="MSPRRPADIRRATEADMPACTIVHNYIETSTVNFRTPEQBPQ
EWTDVLRERYPLVAEVDGEVAGIAYAGPWKARNAYDWTAEISTVYVSPRHQTGL
GSTLYTHLKLEAGFKSVAVIGLNDPDSVRMHEALGYAPRGMRLAAGFKHGNWHD
VGFQWDFSLPVPVRPLPVTET"
16711..17439
/note="contains polyA signal"
source: octopine synthase gene, Agrobacterium tumefaciens,
plasmid pTi15955"
17944..18954
/note="source: Escherichia coli"
/codon_start=1
/transl_table=11
/product="adenyl transferase"
/protein_id="CAA77217.1"
/db_xref="GI:4191252"
/translation="MRSRWGRTLTTERSGNGAVAFMACYDCFGVQSMRASKQQA
RYAGRCMLMSSNDVTQGSRPKTKLMREAVIAEVSTOLSEVVGVIERHLEPTLL
AVHLGSAVDGLKPHSIDLLVTVRLDETTRALINDLETTSASGESEILRAVE
VTIUVHDDIIFWRPKRELQFGQWRNDIAGIFEPATIIDILAILTKARESHVAL
VTPAAEELDFVPEODLFEALNETLWNSPDWAGDNRNVLTLSRIWISAVTGKIA
PKDVAADWAMERLPAQYQVILEARQAYLQGEEDRLASRADOLEEYHVYKGEITKV
GK"
20107..22290
/note="chimeric uidA (gus) gene"
20121..21944
/gene="gus"
20121..21944
/gene="gus"
/note="source: Escherichia coli"
/codon_start=1

Alignment Scores:
Pred. No.: 1.28e-21 Length: 22574
Score: 262.00 Matches: 52
Percent Similarity: 93.4% Conservative: 5
Best Local Similarity: 85.2% Mismatches: 4
Query Match: 85.9% Indels: 0
DB: 8 Gaps: 0

US-10-718-311-15 (1-62) x CVR18556 (1-22574)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 2276 ATGGCTTCTCTGTGATTTCTCTGCACCTGTGTGCTACTCGACCTAAATGTGGCTCAAGCT 2335

Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIleLys 40
Db 2336 AGCATGGTTGCACCTTTTAAATGGTCTTAAGTCTGTGCTCTCTCTCCAGTTTCAAGGAAG 2395

Qy 41 GlnAanLeuAapIleThrSerIleAlaSerAanGlyGlyArgValSerCysMetClnVal 60
Db 2396 CAAACCTTGACATCACCTTCATTCTAGCAATGTGGTGAAGAGTCCAATGCATGCAGGTG 2455

```

```

Qy 61 Trp 61
Db 2456 TGG 2458

RESULT 14
TOBRUBPA
LOCUS TOBRUBPA 979 bp DNA linear PLN 26-MAY-1994
DEFINITION Tobacco ribulose-1,5-bisphosphate carboxylase small subunit gene,
exons 1 and 2, clone TSSU3-8.
ACCESSION M32419.1 GI:170327
VERSION M32419.1
KEYWORDS ribulose 1,5-bisphosphate carboxylase.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 979)
AUTHORS O'Neal J.K., Pokalsky A.R., Kiehne, K.L. and Shewmaker, C.K.
TITLE Isolation of tobacco SSU genes: characterization of a
transcriptionally active pseudogene
JOURNAL Nucleic Acids Res. 15 (21), 8661-8677 (1987)
PUBMED 3684569
COMMENT Original source text: Nicotiana tabacum DNA.
FEATURES
source
1..979
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"
join(584..760,854..>979)
/note="precursor"
/codon_start=1
/product="ribulose-1,5-bisphosphate carboxylase small
subunit"
/db_xref="GI:170328"
/translation="MASSVLSAAVATRSNVAQNMVAPFTGLKSAASFPVSRKONLD
ITSTASNGRVOCMQVWPPINKKYETLSYLPDISQEQLLSEVYLLKNGWPCLEF"
584..757
/sig_peptide
/note="ribulose-1,5-bisphosphate carboxylase small subunit
signal peptide"
join(758..760,854..>979)
/mat_peptide
/product="ribulose-1,5-bisphosphate carboxylase small
subunit"
<584..760
/exon
/note="ribulose-1,5-bisphosphate carboxylase small subunit
precursor"
/number=1
intron
761..853
/note="ribulose-1,5-bisphosphate carboxylase small
subunit"
/exon
854..>979
/note="ribulose-1,5-bisphosphate carboxylase small
subunit; putative"
/number=2

ORIGIN
Alignment Scores:
Pred. No.: 6.9e-23 Length: 979
Score: 258.00 Matches: 52
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.7% Mismatches: 2
Query Match: 84.8% Indels: 0
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOBRUBPA (1-979)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 584 ATGGCTTCTCAGTTCCTTCTCTGCAGCAGTTGCCACCCGACGCAATGTTGCTCAAGCT 643

```

```
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 644 AACATGGTGCACCTTCACTGCGCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 703
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 704 CAAACCTTGACATCACTCCATTCGACGACACGGCGGAGAGTGCATGCATGCAGGTA 763

RESULT 15
NPRBCS8B
LOCUS 2293 bp DNA linear PLN 07-APR-1993
DEFINITION Nicotiana plumbaginifolia rbcS-8B gene for ribulose biphosphate
carboxylase small subunit.
ACCESSION X13711
VERSION 1
KEYWORDS enhancer-like sequence; rbcS gene; ribulose biphosphate
carboxylase.
SOURCE Nicotiana plumbaginifolia (curled-leaved tobacco)
ORGANISM Nicotiana plumbaginifolia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 2293)
Poulsen, C., Fluhr, R., Kauffman, J. M., Boutry, M. and Chua, N. H.
Characterization of an rbcS gene from Nicotiana plumbaginifolia and
expression of an rbcS-CAT chimeric gene in homologous and
heterologous nuclear background
Mol. Gen. Genet. 205, 193-200 (1986)
2 (bases 1 to 1130)
Poulsen, C. and Chua, N. H.
Dissection of 5' upstream sequences for selective expression of the
Nicotiana plumbaginifolia rbcS-8B gene
Mol. Gen. Genet. 214 (1), 16-23 (1988)
3226423

FEATURES
Location/Qualifiers
1..2293
/organism="Nicotiana plumbaginifolia"
/mol_type="genomic DNA"
/db_xref="taxon:4092"
/clone_lib="EMBL4"
1..450
/notes="enhancer-like sequence"
complement(403..410)
/notes="SV40 enhancer core similarity"
435..892
/notes="SV40 enhancer core similarity"
885..892
/notes="SV40 enhancer core similarity"
911..915
CAAT_signal
927..931
CAAT_signal
945..949
TATA_signal
1008..1015
misc_feature
1035
/notes="alternative transcriptional start"
1039..2171
/notes="primary transcript"
1039..1289
/number=1
1049..1051
/notes="alternative transcriptional start"
join(1113..1289,1381..1515,1644..1696,1815..1992)
/codon_start=1
/product="ribulose biphosphate carboxylase"
/protein_id="CAA31994.1"
/db_xref="GI:19706"
/db_xref="GOA:P26573"
/db_xref="UniProtKB/Swiss-Prot:P26573"
/translation="MASSVLSAAVATRSNVAQANVAFPTGLKSNASFPVSRKQNLID
ITSIASNGRVQCMQWPPINKKXETLSLPDLSEQLLSIEYLLKSGWVPCLEFE
TERGFYREHHSPGVDGRYWTMKLPMFGTDTQVLAEEVEAKAYPOAWVRIIG
FDNRVQVCISFIAYKPEGY"
1290..1380
/number=1

misc_feature
misc_feature
misc_feature
misc_feature
CAAT_signal
CAAT_signal
CAAT_signal
TATA_signal
misc_feature
precursor_RNA
exon
misc_feature
CDS
intron
```

```
exon 1381..1515
intron /number=2
exon 1516..1643
intron /number=2
exon 1644..1696
intron /number=3
exon 1697..1814
intron /number=3
exon 1815..2171
misc_feature /number=4
misc_feature 2147..2151
polyA_signal /note="alternative transcriptional stop site"
misc_feature 2154..2159
misc_feature 2167..2170
misc_feature /note="alternative transcriptional stop site"

ORIGIN
Alignment Scores:
Pred. No.: 2 09e-22 Length: 2293
Score: 258.00 Matches: 52
Percent Similarity: 86.7% Conservatives: 6
Best Local Similarity: 96.7% Mismatches: 2
Query Match: 84.6% Indels: 0
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x NPRBCS8B (1-2293)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1113 ATGGCTTCTCAGTTCTTTCTCTGTCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 1172
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 1173 AACATGGTTCACCTTTCCTGCTTAAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 1232
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 1233 CAAACCTTGACATCACTCCATTCGACGACACGGCGGAGAGTGCATGCATGCAGGTA 1292
```

Search completed: July 11, 2006, 21:35:35  
Job time : 1360.71 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 20:58:40 ; Search time 36.8997 Seconds

(without alignments)  
3276.788 Million cell updates/sec

Title: US-10-718-311-15

Perfect score: 305

Sequence: 1 MASSVISAAVATRSNTQA.....LDITSIASNGRVCQMWH 62

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US10718311/runat\_11072006.111535\_5434/app\_query.fasta.1  
-DB=Published Applications NA New -OFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abss06p  
-USER=US10718311@CGN 1.1.251 @runat\_11072006.111535\_5434 -NCPU=6 -ICPU=3  
-NO MNAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.\*

1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269	88.2	762	8	US-11-247-587-125 Sequence 125, App
2	269	88.2	781	8	US-11-247-587-124 Sequence 124, App
3	250.5	82.1	698	8	US-11-247-587-128 Sequence 128, App
4	244.5	80.2	683	8	US-11-247-587-126 Sequence 126, App
5	242.5	79.5	700	8	US-11-247-587-127 Sequence 127, App
6	231	75.7	583	7	US-11-218-305-17631 Sequence 17631, A
7	227.5	74.6	727	8	US-11-247-587-129 Sequence 129, App
8	225	73.8	743	6	US-10-953-349-8894 Sequence 8894, App

9	224	73.4	658	7	US-11-218-305-17633 Sequence 17633, A
10	221	72.5	718	6	US-10-953-349-10521 Sequence 10521, A
11	215	70.5	640	7	US-11-218-305-17634 Sequence 17634, A
12	201	65.9	799	6	US-10-953-349-9569 Sequence 9569, App
13	201	65.9	828	7	US-11-218-305-17632 Sequence 17632, A
14	194	63.6	841	8	US-11-216-545-4173 Sequence 4173, App
15	194	63.6	1023	8	US-11-216-545-4174 Sequence 4174, App
16	183	60.0	1237	8	US-11-216-545-8381 Sequence 8381, App
17	131	43.0	947	6	US-10-449-902-24132 Sequence 24132, A
18	129	42.3	627	8	US-11-216-545-6801 Sequence 6801, App
19	129	42.3	882	8	US-11-266-748A-366894 Sequence 366894, App
20	129	42.3	882	8	US-11-266-748A-388679 Sequence 388679, App
21	129	42.3	882	8	US-11-266-748A-450273 Sequence 450273, App
22	125	41.0	845	6	US-10-449-902-19918 Sequence 19918, A
23	125	41.0	1999	6	US-10-449-902-17846 Sequence 17846, A
24	122	40.0	830	6	US-10-449-902-4174 Sequence 4174, App
25	120	39.3	899	6	US-10-449-902-18147 Sequence 18147, A
26	118	38.7	2409	6	US-10-449-902-20231 Sequence 20231, A
27	110	36.1	416	6	US-10-523-290-17 Sequence 17, Appl
28	110	36.1	5600	6	US-10-525-318-13 Sequence 13, Appl
29	110	36.1	7000	6	US-10-525-318-8 Sequence 8, Appl
30	72.5	23.8	3004	6	US-10-449-902-14773 Sequence 14773, A
31	69.5	22.8	964	8	US-11-266-748A-414508 Sequence 414508, App
32	66.5	21.8	1875	8	US-11-217-529-76380 Sequence 76380, A
33	66.5	21.8	4223	7	US-11-218-305-16072 Sequence 16072, A
34	65.5	21.5	1613	7	US-11-218-305-13081 Sequence 13081, A
35	64.5	21.1	1679	6	US-10-449-902-28323 Sequence 28323, A
36	64	21.0	1493	8	US-11-266-748A-185234 Sequence 185234, App
37	64	21.0	1493	8	US-11-266-748A-192750 Sequence 192750, App
38	63	20.7	1373	8	US-11-216-545-6828 Sequence 6828, App
39	63	20.7	1510	6	US-10-953-349-14286 Sequence 14286, A
40	63	20.7	2519	6	US-10-449-902-7755 Sequence 7755, App
41	63	20.7	3811	7	US-11-218-305-9560 Sequence 9560, App
42	63	20.7	3951	8	US-11-121-154-85 Sequence 85, Appl
43	62.5	20.5	1337	8	US-11-266-748A-368662 Sequence 368662, App
44	62.5	20.5	1337	8	US-11-266-748A-452041 Sequence 452041, App
45	62.5	20.5	1406	6	US-10-449-902-20240 Sequence 20240, A

#### ALIGNMENTS

RESULT 1  
US-11-247-587-125  
; Sequence 125, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 125  
; LENGTH: 762  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana  
US-11-247-587-125

Alignment Scores:  
Pred. No.: 2.86e-31 Length: 762  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6

Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 8 Gaps: 0

US-10-718-311-15 (1-62) x US-11-247-587-125 (1-762)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 27 ATGGCTTCTCAGTTCTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 86  
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyLys 40  
 Db 87 AACATGGTTGCACCTTCACTGGTCTTAAGTCAGCTGCCTCGTTCCCTGTTTCAAGGAAG 146  
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 147 CAAACACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCATATGCATGCAGGTG 206  
 Qy 61 Trp 61  
 Db 207 TGG 209

#### RESULT 2

US-11-247-587-124  
 ; Sequence 124, Application US/11247587  
 ; Publication No. US20060130176A1

#### GENERAL INFORMATION:

; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; PRIOR FILING DATE: 2005-10-12  
 ; PRIOR APPLICATION NUMBER: US 60/671,089  
 ; PRIOR FILING DATE: 2005-04-14  
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
 ; PRIOR FILING DATE: 2004-10-12  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
 ; SEQ ID NO 124

LENGTH: 781

TYPE: DNA

ORGANISM: Nicotiana benthamiana

US-11-247-587-124

Alignment Scores:  
 Pred. No.: 2,978-31 Length: 781  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 8 Gaps: 0

US-10-718-311-15 (1-62) x US-11-247-587-124 (1-781)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 39 ATGGCTTCTCAGTTCTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 98  
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyLys 40  
 Db 99 AACATGGTTGCACCTTCACTGGTCTTAAGTCAGCTGCCTCGTTCCCTGTTTCAAGGAAG 158  
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 159 CAAACACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCATATGCATGCAGGTG 218  
 Qy 61 Trp 61

Db 219 TGG 221

#### RESULT 3

US-11-247-587-128  
 ; Sequence 128, Application US/11247587  
 ; Publication No. US20060130176A1

#### GENERAL INFORMATION:

; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; PRIOR FILING DATE: 2005-10-12  
 ; PRIOR APPLICATION NUMBER: US 60/671,089  
 ; PRIOR FILING DATE: 2005-04-14  
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
 ; PRIOR FILING DATE: 2004-10-12  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
 ; SEQ ID NO 128

LENGTH: 698

TYPE: DNA

ORGANISM: Nicotiana benthamiana

US-11-247-587-128

Alignment Scores:  
 Pred. No.: 1,598-28 Length: 698  
 Score: 250.50 Matches: 52  
 Percent Similarity: 91.9% Conservative: 5  
 Best Local Similarity: 83.9% Mismatches: 4  
 Query Match: 82.1% Indels: 1  
 DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-128 (1-698)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGln 19  
 Db 20 ATGGCTTCTCAGTTCTTCTCAGCAGTCTGCCACCGCAGCAATGTTGCTCAAG 79  
 Qy 20 AlaSerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIly 39  
 Db 80 GCCAGCATGGTTGCACCTTCACTGGCTCAATCTGCTTCTCTCTCTGTTACCAGA 139  
 Qy 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59  
 Db 140 AACAAAAACCTTGACATTACATTCATTGCTAGCAATGTTGGAAGAGTCCAATGCATGCAG 199  
 Qy 60 ValTrp 61  
 Db 200 GTGTGG 205

#### RESULT 4

US-11-247-587-126  
 ; Sequence 126, Application US/11247587  
 ; Publication No. US20060130176A1

#### GENERAL INFORMATION:

; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; PRIOR FILING DATE: 2005-10-12



```
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 126
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-126

Alignment Scores:
Pred. No.: 1,24e-27 Length: 683
Score: 244.50 Matches: 50
Percent Similarity: 91.9% Conservatives: 7
Best Local Similarity: 80.6% Mismatches: 4
Query Match: 80.2% Indels: 1
DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-126 (1-683)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
Db 37 ATGGCTTCTCAGTTATGCTTCCTCAGCAGCTGCTGTGGACCGGCCCAATGCTGCTCAA 96
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
Db 97 GCCAACATGTTGCACCCCTTCACTGGCCTCAAGTCGGCTCCTCCTCTCCCTGTTACCAGG 156
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 157 AAACAAAACCTTGACATTACTTCCATTGCTAGCAATGGTGAAGAGTTCAATGCATGCAG 216
QY 60 ValTrp 61
Db 217 GTGTGG 222

RESULT 5
US-11-247-587-127
; Sequence 127, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; PRIOR FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-127

Alignment Scores:
Pred. No.: 2,58e-27 Length: 700
Score: 242.50 Matches: 50
Percent Similarity: 90.3% Conservatives: 6
Best Local Similarity: 80.6% Mismatches: 5
Query Match: 79.5% Indels: 1
DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-127 (1-700)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
Db 2 ATGGCTTCTCAGTTATGCTTCCTCAGCTGCGCTGTTGCCACCGGCCCAATGCTGCTCAA 61
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
Db 62 GCCAGTATGGTTGCACCTTTCACCTGGCCTCAAGTCGGCAACCTCCTTCCCTGTTCCAGA 121
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 122 AAACAAAACCTTGACATTACTTCCATTGCTAGCAACCGCGGAGAGAGTTCAATGCATGCAG 181
QY 60 ValTrp 61
Db 182 GTGTGG 187

RESULT 6
US-11-218-305-17631
; Sequence 17631, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Mc Laird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17631
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17631

Alignment Scores:
Pred. No.: 1,09e-25 Length: 583
Score: 231.00 Matches: 48
Percent Similarity: 88.5% Conservatives: 6
Best Local Similarity: 78.7% Mismatches: 5
Query Match: 75.7% Indels: 2
DB: 7 Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17631 (1-583)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTCTCAGTTATGCTTCCTCAGCTGCGCTGTTGGCTACC-----TCACCGGCTCAAGCC 85
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTTGCTCCATTACCGGCTTGAAGTCATCGCTGCATTCCTCCAGTCACCCGCAAA 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 AGCAACACTGATATTACTTCCATTGCAAGCAACGGAGGAAGAGTTAGTGCATGAAGGTG 206
QY 61 Trp 61
Db 207 TGG 209

RESULT 7
US-11-247-587-129
; Sequence 129, Application US/11247587
; Publication No. US20060130176A1
```

;

```
US-10-953-349-10521
; Sequence 10521, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10521
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10521

Alignment Scores:
Pred. No.: 4,79e-24 Length: 718
Score: 221.00 Matches: 46
Percent Similarity: 86.9% Conservative: 7
Best Local Similarity: 75.4% Mismatches: 6
Query Match: 72.5% Indels: 2
DB: Gaps: 1

US-10-718-311-15 (1-62) x US-10-953-349-10521 (1-718)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaAlaThrArgSerAsnValThrGlnAla 20
Db 32 ATGGCTTCTCTATGCTCTCTCTCGCGCTGTGTTACC-----TCCCGGCTCAAGCC 85
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 86 ACCATGGTCGCTCCATTCACTGGTTTGAAGTCATCGCTTCTTTCCGGTCAACCCGCAAG 145
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 146 GCCAACACGACATTACTTCCATCAACGAATGGGGGAAGAGTTAGCTGCATGAAGGTG 205
QY 61 Trp 61
Db 206 TGG 208

RESULT 11
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634

Alignment Scores:
Pred. No.: 5,9e-21 Length: 799
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: Gaps: 1

US-10-718-311-15 (1-62) x US-10-953-349-9569 (1-799)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaAlaThrArgSerAsnValThrGlnAla 20
Db 47 ATGGCTTCTCTATGCTCTCTCTCGCGCTACTAGTT-----GCCTCTCCGGCTCAAGCC 100
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 101 ACTATGGTCGCTCCITTCACGAGACTTAAGTCTCTCGCTGCCTTCCAGCCACCCGCAAG 160
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 161 GCTAACACGACATTACTTCCATCAACGAATGGGGGAAGAGTTAACTGCATGCAGGTG 220

RESULT 13
US-11-218-305-17632
; Sequence 17632, Application US/11218305
```

```
US-10-953-349-9569
; Sequence 9569, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9569
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9569

Alignment Scores:
Pred. No.: 5,9e-21 Length: 799
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17634 (1-640)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTACTCTATGCTCTCTCTCGCGCTGTGTTACC-----TCCCGGCTCAAGCC 86
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTCGCTCCATTCAACGAGTCTTCTCTGCATTCACCCGCAAG 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 GCCAACACGACATTACTTCCATCGTTAGCAACGAGGAAGAGTTAGCTGCATGCAGGTG 206
QY 61 Trp 61
Db 207 TGG 209

RESULT 12
US-10-953-349-9569
; Sequence 9569, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9569
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9569

Alignment Scores:
Pred. No.: 5,9e-21 Length: 799
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17634 (1-640)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTACTCTATGCTCTCTCTCGCGCTGTGTTACC-----TCCCGGCTCAAGCC 86
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTCGCTCCATTCAACGAGTCTTCTCTGCATTCACCCGCAAG 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 GCCAACACGACATTACTTCCATCGTTAGCAACGAGGAAGAGTTAGCTGCATGCAGGTG 206
QY 61 Trp 61
Db 207 TGG 209

RESULT 13
US-11-218-305-17632
; Sequence 17632, Application US/11218305
```

```

Score: 194.00      Matches: 45
Percent Similarity: 77.0%      Conservative: 2
Best Local Similarity: 73.8%      Mismatches: 12
Query Match: 63.6%      Indels: 2
DB: 8      Gaps: 2

US-10-718-311-15 (1-62) x US-11-216-545-4173 (1-841)

Qy 1 MetLaserSerValIleSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
    |||||:|||||
Db 37 ATGGCTTCCTCAATGATCTCTCCACAGCTGTACCACC--GTCAACCGTGCCTGGTCC 93
    |||||:|||||
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
    |||||:|||||
Db 94 GGCATGGTGTCTCAATTACCGGCTCAATCCATGGCTGGCTTCCCC---ACGAGGAAG 150
    |||||:|||||
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
    |||||:|||||
Db 151 ACCAACAAATGACATTACCTCCATTGCTAGCAACGGTGAAGAGTAGTCAATGCGATGCAGGTG 210
    |||||:|||||
Qy 61 Trp 61
    |||
Db 211 TGG 213

RESULT 15
US-11-216-545-4174
; Sequence 4174, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Ge
; TITLE OF INVENTION: Soybeans.
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4174
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-4174

Alignment Scores:
Pred No.: 9,62e-20      Length: 1023
Score: 194.00      Matches: 45
Percent Similarity: 77.0%      Conservative: 2
Best Local Similarity: 73.8%      Mismatches: 12
Query Match: 63.6%      Indels: 2
DB: 8      Gaps: 2

US-10-718-311-15 (1-62) x US-11-216-545-4174 (1-1023)

Qy 1 MetLaserSerValIleSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
    |||||:|||||
Db 228 ATGGCTTCCTCAATGATCTCTCCACAGCTGTACCACCT--GTCAACCGTGCCTGGTCC 284
    |||||:|||||
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
    |||||:|||||
Db 285 GGCATGGTGTCTCAATTACCTGGCTCAAGTCCATGGCTGGCTTCCCC---ACCAGGAAG 341
    |||||:|||||
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
    |||||:|||||
Db 342 ACCAACAAATGACATTACCTCCATTGCTAGCAACGGTGAAGAGTAGTCAATGCGATGCAGGTG 401
    |||||:|||||

```

Qy 61 Trp 61  
|||  
Db 402 TGG 404

Search completed: July 11, 2006, 21:05:59  
Job time : 37.8997 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 20:23:26 ; Search time 452.45 Seconds  
(without alignments)  
2525.695 Million cell updates/sec

Title: US-10-718-311-15  
Perfect score: 305  
Sequence: 1 MASSVISAATVRSNVTAQ.....LDITSIANGRVSCHQVWH 62

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delext 7.0	

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abes/ABSSWEB spool/US10718311/runat 11072006 111531 5371/app\_query.fasta\_1  
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss08  
-USER=US10718311 @CN 1 1 2326 @runat 11072006 111531 5371 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

- 1: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 3: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 6: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 7: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 8: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 9: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 10: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 11: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 12: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 13: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 14: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*
- 15: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*
- 16: /EMC Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	305	100.0	684	3	US-09-855-341-7 Sequence 7, Appli

2	305	100.0	684	3	US-09-896-866B-16
3	305	100.0	684	7	US-10-359-369-41
4	305	100.0	684	8	US-10-699-050-16
5	305	100.0	684	8	US-10-718-311-7
6	305	100.0	684	9	US-10-462-162-33
7	269	88.2	204	3	US-09-839-477-3
8	269	88.2	204	8	US-10-758-064-3
9	269	88.2	297	10	US-10-487-901-7209
10	269	88.2	377	10	US-10-487-901-3504
11	269	88.2	489	10	US-10-487-901-3503
12	269	88.2	504	10	US-10-487-901-7207
13	269	88.2	718	10	US-10-487-901-5765
14	269	88.2	736	10	US-10-487-901-1704
15	269	88.2	736	10	US-10-487-901-5043
16	269	88.2	736	10	US-10-487-901-7208
17	253.5	83.1	704	9	US-10-425-115-85509
18	250.5	82.1	683	10	US-10-487-901-1713
19	250.5	82.1	737	10	US-10-487-901-5045
20	250.5	82.1	737	10	US-10-487-901-5766
21	244.5	80.2	738	10	US-10-487-901-1716
22	244.5	80.2	753	10	US-10-487-901-5764
23	244.5	80.2	754	10	US-10-487-901-5028
24	244	80.0	174	3	US-09-854-286-15
25	242.5	79.5	714	10	US-10-487-901-1703
26	242.5	79.5	736	10	US-10-487-901-5044
27	235	77.0	169	6	US-10-165-420-3
28	232.5	76.2	654	10	US-10-487-901-5042
29	225	73.8	385	3	US-09-732-627A-1176
30	225	73.8	724	8	US-09-910-664-55
31	225	73.8	724	8	US-10-333-184-54
32	225	73.8	726	8	US-10-333-184-202
33	225	73.8	771	3	US-09-770-445-918
34	225	73.8	1205	3	US-09-770-445-41
35	221	72.5	363	3	US-09-770-791-468
36	221	72.5	711	8	US-09-910-664-32
37	221	72.5	711	8	US-10-333-184-31
38	221	72.5	739	8	US-10-333-184-182
39	221	72.5	1036	9	US-10-739-930-2595
40	217	71.1	767	3	US-09-770-445-931
41	217	71.1	780	3	US-09-770-445-880
42	214	70.2	362	3	US-09-770-791-463
43	214	70.2	392	3	US-09-770-423-971
44	213.5	70.0	424	10	US-10-487-901-1710
45	213.5	70.0	424	10	US-10-487-901-5032

ALIGNMENTS

RESULT 1  
US-09-855-341-7  
; Sequence 7, Application US/09855341  
; Patent No. US20020002715A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; TITLE OF INVENTION: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL  
US-09-855-341-7

Alignment Scores: 8.14e-36 Length: 684  
Pred. No.: 1

```
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-15 (1-62) x US-09-855-341-7 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTGCTATTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGACCTTCACTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 2
; Sequence 16, Application US/09896866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose-Malate Sinapoyltransferase Form Malate Conjugates
; FILE REFERENCE: BCI034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; NAME/KEY: misc_feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-09-896-866B-16

Alignment Scores:
Pred. No.: 8 14e-36 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-15 (1-62) x US-09-896-866B-16 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTGCTATTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGACCTTCACTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHis 62
```

```
Db 181 TGGCAT 186

RESULT 3
US-10-359-369-41
; Sequence 41, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew
; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CL1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 41
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
; NAME/KEY: CDS
; LOCATION: (1)..(684)
US-10-359-369-41

Alignment Scores:
Pred. No.: 8 14e-36 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x US-10-359-369-41 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTGCTATTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGACCTTCACTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 4
US-10-699-050-16
; Sequence 16, Application US/10699050
; Publication No. US20040142437A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose-Malate Sinapoyltransferase Form Malate Conjugates
; FILE REFERENCE: BCI034 US NA
; CURRENT APPLICATION NUMBER: US/10/699,050
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
```



```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-10-699-050-16

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-718-311-15 (1-62) x US-10-699-050-16 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGTCAGCAATGGTGGAGAGTTAGTGCATGCAGGTG 180
QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 5
US-10-718-311-7
; Sequence 7, Application US/10718311
; Publication No. US20040143867A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUIT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/10/718,311
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/855,341
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-10-718-311-7

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-718-311-15 (1-62) x US-10-718-311-7 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGTCAGCAATGGTGGAGAGTTAGTGCATGCAGGTG 180
QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 6
US-10-718-311-15 (1-62) x US-10-462-162-33 (1-684)
; Sequence 33, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-33

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             9          Gaps:      0

US-10-718-311-15 (1-62) x US-10-462-162-33 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGTCAGCAATGGTGGAGAGTTAGTGCATGCAGGTG 180
QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 7
US-09-839-477-3
; Sequence 3, Application US/09839477
; Publication No. US20030167523A1
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
```

```
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGTCAGCAATGGTGGAGAGTTAGTGCATGCAGGTG 180
QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 6
US-10-462-162-33
; Sequence 33, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-33

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             9          Gaps:      0

US-10-718-311-15 (1-62) x US-10-462-162-33 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGTCAGCAATGGTGGAGAGTTAGTGCATGCAGGTG 180
QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 7
US-09-839-477-3
; Sequence 3, Application US/09839477
; Publication No. US20030167523A1
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
```

```
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(204)
US-09-839-477-3

Alignment Scores:
Pred. No.: 4,62e-31 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-15 (1-62) x US-09-839-477-3 (1-204)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTTCTTCTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCACTGGCCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 180

Qy 61 Trp 61
Db 181 TGG 183

RESULT 8
US-10-758-064-3
; Sequence 3, Application US/10758064
; Publication No. US20040128718A1
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/10/758,064
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/839,477
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(204)
US-10-758-064-3

Alignment Scores:
Pred. No.: 4,62e-31 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 8 Gaps: 0
```

```
US-10-718-311-15 (1-62) x US-10-758-064-3 (1-204)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTTCTTCTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCACTGGCCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 180

Qy 61 Trp 61
Db 181 TGG 183

RESULT 9
US-10-487-901-7209
; Sequence 7209, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McGreery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7209
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-7209

Alignment Scores:
Pred. No.: 7,79e-31 Length: 297
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-7209 (1-297)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 97 ATGGCTTCTCAGTTCTTCTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 156

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 157 AACATGGTTGCACCTTCACTGGCCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 216

Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 217 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 276

Qy 61 Trp 61
Db 181 TGG 183
```

```
Db 277 TGG 279
RESULT 10
US-10-487-901-3504
; Sequence 3504, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; CURRENT APPLICATION NUMBER: US/10/487,901
; FILE REFERENCE: DOW-08552
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3504
; LENGTH: 377
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-3504
Alignment Scores:
Pred. No.: 1.08e-30 Length: 377
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0
US-10-718-311-15 (1-62) x US-10-487-901-3504 (1-377)
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; CURRENT APPLICATION NUMBER: US/10/487,901
; FILE REFERENCE: DOW-08552
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3504
; LENGTH: 377
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-3504
Alignment Scores:
Pred. No.: 1.08e-30 Length: 377
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0
US-10-718-311-15 (1-62) x US-10-487-901-3504 (1-377)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 68 ATGGCTTCCTCAGTTCTTCTCCAGCAGCAGTTCACCCGCGAGCAATGTTGCTCAAGCT 127
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 128 AACATGGTTGCACCTTTCACAGGCTCTTAAGTCTGCTGCTCATTCCCTGTTTCAAGAAAG 187
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 188 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 247
QY 61 TTP 61
Db 248 TGG 250
RESULT 11
US-10-487-901-3503
; Sequence 3503, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; CURRENT APPLICATION NUMBER: US/10/487,901
; FILE REFERENCE: DOW-08552
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3503
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-3503
Alignment Scores:
Pred. No.: 1.56e-30 Length: 489
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0
US-10-718-311-15 (1-62) x US-10-487-901-3503 (1-489)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 79 ATGGCTTCCTCAGTTCTTCTCCAGCAGCAGTTCACCCGCGAGCAATGTTGCTCAAGCT 138
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 139 AACATGGTTGCACCTTTCACAGGCTCTTAAGTCTGCTGCTCATTCCCTGTTTCAAGAAAG 198
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 199 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 258
QY 61 TTP 61
Db 259 TGG 261
RESULT 12
US-10-487-901-7207
; Sequence 7207, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; CURRENT APPLICATION NUMBER: US/10/487,901
; FILE REFERENCE: DOW-08552
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7207
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-7207
Alignment Scores:
Pred. No.: 1.62e-30 Length: 504
```

Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-7207 (1-504)

Qy 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 78 ATGGCTTCTCAGTCTTCTTCTCAGCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 137  
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40  
Db 138 AACATGGTTGCACCTTTCACAGTCTTAAGTCTGTGCTCATTCCTGTTTCAAGAAG 197  
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60  
Db 198 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATGCAGGTG 257  
Qy 61 Trp 61  
Db 258 TGG 260

## RESULT 13

US-10-487-901-5765  
; Sequence 5765, Application US/10487901  
; Publication No. US20050091708A1  
; GENERAL INFORMATION:  
; APPLICANT: Oreido, Jeremiah Vincent  
; APPLICANT: McCrery, David  
; APPLICANT: Pell, Randy  
; APPLICANT: Miller, Barbara  
; APPLICANT: Weglarz, Thaddeus  
; APPLICANT: Gachotte, Daniel  
; APPLICANT: Blakeslee, Beth  
; APPLICANT: Larrinua, Ignacio  
; APPLICANT: Reddy, Avutu  
; APPLICANT: Shukla, Vipula  
; APPLICANT: Crosley, Rodney  
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri  
; FILE REFERENCE: DOW-08552  
; CURRENT APPLICATION NUMBER: US/10/487,901  
; CURRENT FILING DATE: 2004-02-26  
; NUMBER OF SEQ ID NOS: 7560  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5765  
; LENGTH: 718  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic

## US-10-487-901-5765

Alignment Scores:  
Pred. No.: 2,666-30 Length: 718  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-5765 (1-718)

Qy 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 45 ATGGCTTCTCAGTCTTCTTCTCAGCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 104  
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40  
Db 105 AACATGGTTGCACCTTTCACAGTCTTAAGTCAAGTGTGCTGCTCCCTGTTTCAAGAAG 164  
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60

Db 165 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATGCAGGTG 224  
Qy 61 Trp 61  
Db 225 TGG 227

## RESULT 14

US-10-487-901-1704  
; Sequence 1704, Application US/10487901  
; Publication No. US20050091708A1  
; GENERAL INFORMATION:  
; APPLICANT: Oreido, Jeremiah Vincent  
; APPLICANT: McCrery, David  
; APPLICANT: Pell, Randy  
; APPLICANT: Miller, Barbara  
; APPLICANT: Weglarz, Thaddeus  
; APPLICANT: Gachotte, Daniel  
; APPLICANT: Blakeslee, Beth  
; APPLICANT: Larrinua, Ignacio  
; APPLICANT: Reddy, Avutu  
; APPLICANT: Shukla, Vipula  
; APPLICANT: Crosley, Rodney  
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri  
; FILE REFERENCE: DOW-08552  
; CURRENT APPLICATION NUMBER: US/10/487,901  
; CURRENT FILING DATE: 2004-02-26  
; NUMBER OF SEQ ID NOS: 7560  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1704  
; LENGTH: 736  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic

## US-10-487-901-1704

Alignment Scores:  
Pred. No.: 2,756-30 Length: 736  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-1704 (1-736)

Qy 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 63 ATGGCTTCTCAGTCTTCTTCTCAGCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 122  
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40  
Db 123 AACATGGTTGCACCTTTCACAGTCTTAAGTCAAGTGTGCTGCTCCCTGTTTCAAGAAG 182  
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60  
Db 183 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATGCAGGTG 242  
Qy 61 Trp 61  
Db 243 TGG 245

## RESULT 15

US-10-487-901-5043  
; Sequence 5043, Application US/10487901  
; Publication No. US20050091708A1  
; GENERAL INFORMATION:  
; APPLICANT: Oreido, Jeremiah Vincent  
; APPLICANT: McCrery, David  
; APPLICANT: Pell, Randy  
; APPLICANT: Miller, Barbara  
; APPLICANT: Weglarz, Thaddeus

; APPLICANT: Gachotte, Daniel  
; APPLICANT: Blakeslee, Beth  
; APPLICANT: Larrinua, Ignacio  
; APPLICANT: Reddy, Avutu  
; APPLICANT: Shukla, Vipula  
; APPLICANT: Croseley, Rodney  
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri  
; FILE REFERENCE: DOW-08552  
; CURRENT APPLICATION NUMBER: US/10/487,901  
; CURRENT FILING DATE: 2004-02-26  
; NUMBER OF SEQ ID NOS: 7560  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5043  
; LENGTH: 736  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-487-901-5043

Alignment Scores:  
Pred. No.: 2.75e-30 Length: 736  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-5043 (1-736)

Qy	1	MetAlaSerSerValIleSerSerAlaAlaValAlaAlaThrArgSerAsnValThrGlnAla	20
Db	63	ATGGCTTCCTCAGTTCTTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT	122
Qy	21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40
Db	123	AACATGGTTGCACCTTTCACCTGGTCTTAAGTCAGTCGCTGCTTCCTGTTTCAAGGAAG	182
Qy	41	GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60
Db	183	CAAAACCTTGACATCACTTCATTGCCAGCAACGCGGAAGAGTGCAATGCATGCAGGTG	242
Qy	61	Trp	61
Db	243	TGG	245

Search completed: July 11, 2006, 22:12:04  
Job time : 452.45 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 20:16:09 ; Search time 54.7059 Seconds  
(without alignments)  
3180.885 Million cell updates/sec

Title: US-10-718-311-15  
Perfect score: 305  
Sequence: 1 MASSVSSNAVATRSNTQA.....LDITSIASNGRVSCMQVWH 62

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 segs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10718311/runat 11072006 111527 5281/app query.fasta\_1  
-DB=Issued Patents NA -QFIX=fastp -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs06p -USER=US10718311 @CN 1 1 381 @runat 11072006 111527 5281  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*\*  
1: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*\*  
2: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*\*  
3: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*\*  
4: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*\*  
5: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*\*  
6: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*\*  
7: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq:\*\*  
8: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/PP COMB.seq:\*\*  
9: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*\*  
10: /EMC\_Celeerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	305	100.0	684 3 US-09-896-866B-16 Sequence 16, Appl
2	305	100.0	684 3 US-09-855-341-7 Sequence 7, Appl
3	269	88.2	204 3 US-09-839-477-3 Sequence 3, Appl
4	269	88.2	499 3 US-09-495-797-41 Sequence 41, Appl
5	254	83.3	177 2 US-08-095-726-12 Sequence 12, Appl
6	254	83.3	177 2 US-08-096-043-12 Sequence 12, Appl
7	254	83.3	177 2 US-08-093-577-8 Sequence 8, Appl
8	254	83.3	177 2 US-08-331-004A-6 Sequence 6, Appl

9	254	83.3	177	2	US-08-096-623A-12	Sequence 12, Appl
10	254	83.3	177	7	PCT-US95-13937A-6	Sequence 6, Appl
11	249	81.6	1442	2	US-08-152-483B-8	Sequence 8, Appl
12	201	65.9	264	3	US-09-441-340-9	Sequence 9, Appl
13	201	65.9	268	3	US-09-186-002-5	Sequence 5, Appl
14	201	65.9	279	2	US-08-191-139-9	Sequence 9, Appl
15	201	65.9	279	2	US-08-484-274A-9	Sequence 9, Appl
16	201	65.9	279	9	US-09-612-404-9	Sequence 9, Appl
17	201	65.9	355	2	US-08-090-523-5	Sequence 5, Appl
18	201	65.9	355	2	US-08-398-627-5	Sequence 5, Appl
19	201	65.9	355	2	US-08-406-858-5	Sequence 5, Appl
20	201	65.9	355	2	US-08-476-519-7	Sequence 7, Appl
21	201	65.9	355	3	US-08-120-703A-5	Sequence 5, Appl
22	201	65.9	355	3	US-08-399-023-5	Sequence 5, Appl
23	201	65.9	355	7	PCT-US91-04036-5	Sequence 5, Appl
24	201	65.9	355	7	PCT-US94-05275-5	Sequence 5, Appl
25	201	65.9	355	7	PCT-US95-09323-7	Sequence 7, Appl
26	201	65.9	509	3	US-09-011-151-1	Sequence 1, Appl
27	201	65.9	696	3	US-09-441-340-11	Sequence 11, Appl
28	201	65.9	1178	3	US-09-434-039A-36	Sequence 36, Appl
29	201	65.9	1224	3	US-09-434-039A-34	Sequence 34, Appl
30	201	65.9	1630	3	US-09-441-340-23	Sequence 23, Appl
31	201	65.9	8012	3	US-09-182-117-1	Sequence 1, Appl
32	201	65.9	8012	3	US-09-434-039A-1	Sequence 1, Appl
33	201	65.9	8418	3	US-09-182-117-5	Sequence 5, Appl
34	201	65.9	8418	3	US-09-434-039A-5	Sequence 5, Appl
35	201	65.9	8798	3	US-09-182-117-4	Sequence 4, Appl
36	201	65.9	8798	3	US-09-434-039A-4	Sequence 4, Appl
37	201	65.9	10339	3	US-09-186-002-13	Sequence 13, Appl
38	201	65.9	10846	3	US-09-098-219B-5	Sequence 5, Appl
39	201	65.9	10846	3	US-10-164-204-5	Sequence 5, Appl
40	201	65.9	10846	3	US-09-923-109-5	Sequence 5, Appl
41	194	63.6	281	2	US-08-469-202-15	Sequence 15, Appl
42	194	63.6	281	2	US-08-469-202-17	Sequence 17, Appl
43	194	63.6	281	2	US-08-469-202-19	Sequence 19, Appl
44	194	63.6	281	2	US-08-484-434C-15	Sequence 15, Appl
45	194	63.6	281	2	US-08-484-434C-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-09-896-866B-16  
; Sequence 16, Application US/09896866B  
; Patent No. 6642036  
; GENERAL INFORMATION:  
; APPLICANT: Flint, Dennis  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate  
; FILE REFERENCE: BC1034 US NA  
; CURRENT APPLICATION NUMBER: US/09/896,866B  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/216,615  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei  
US-09-896-866B-16

Alignment Scores:  
Pred. No.: 7.84e-35 Length: 684  
Score: 305.00 Matches: 62  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0

```
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-896-866B-16 (1-684)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTGGCACTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTAGCTGCATGCGAGTG 180
QY 61 TrpHis 62
DB 181 TGGCAT 186
RESULT 2
US-09-855-341-7
; Sequence 7, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; FILE REFERENCE: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-09-855-341-7
Alignment Scores:
Pred. No.: 7,84e-35 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-855-341-7 (1-684)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTGGCACTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTAGCTGCATGCGAGTG 180
QY 61 TrpHis 62
DB 181 TGGCAT 186
RESULT 3
US-09-839-477-3
; Sequence 3, Application US/09839477
; Patent No. 6723895
```

```
; GENERAL INFORMATION:
; APPLICANT: Shorrosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: COA-CARBOXYLASE
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-09-839-477-3
Alignment Scores:
Pred. No.: 2,32e-30 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-839-477-3 (1-204)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AACATGGTGGCACTTTCACCTTTCAGCTTAAGTCAGCTGCTCATTCCCTGTTTCAAGGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTAGCTGCATGCGAGTG 180
QY 61 Trp 61
DB 181 TGG 183
RESULT 4
US-09-495-797-41
; Sequence 41, Application US/09495797
; Patent No. 6369296
; GENERAL INFORMATION:
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Mewburn 43,047
; CURRENT APPLICATION NUMBER: US/09/495,797
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence
US-09-495-797-41
Alignment Scores:
Pred. No.: 8,49e-30 Length: 499
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
```







Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAG 177

RESULT 9

US-08-096-623A-12

; Sequence 12, Application US/08096623A

; Patent No. 5684238

; GENERAL INFORMATION:

; APPLICANT: Ausich, Rodney L.

; APPLICANT: Brinkhaus, Friedhelm L.

; APPLICANT: Mukharji, Indrani

; APPLICANT: Proffitt, John H.

; APPLICANT: Yarger, James G.

; APPLICANT: Yen, Hwei-Che B.

; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and

; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts

; NUMBER OF SEQUENCES: 104

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Welsh & Katz, Ltd.

; STREET: 120 S. Riverside Plaza, 22nd Floor

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/096,623A

; FILING DATE: 22-JUL-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/805,061

; FILING DATE: 09-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/662,921

; FILING DATE: 28-FEB-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/562,674

; FILING DATE: 03-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/525,551

; FILING DATE: 18-MAY-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/487,613

; FILING DATE: 02-MAR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Gamson, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER: AMO-006.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 655-1500

; TELEFAX: (312) 655-1501

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 177 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; POSITION IN GENOME:

; MAP POSITION: 1 to 177

; UNITS: bp

US-08-096-623A-12

Alignment Scores:

Pred. No.: 2,87e-28 Length: 177

Score: 254.00 Matches: 51

Percent Similarity: 96.6% Conservative: 2

Best Local Similarity: 86.4% Mismatches: 0

Query Match: 83.3% Indels: 0

DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x US-08-096-623A-12 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCAGTTCTTTCTCTGAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 60

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AACATGGTGGCGCCTTTCACTGGCCTTAAGTCAGTCGCTCATTCCTCTTTTCAAGGAAG 120

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59

Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAG 177

RESULT 10

PCT-US95-13937A-6

; Sequence 6, Application PC/TUS9513937A

; GENERAL INFORMATION:

; APPLICANT: Hauptmann, Randal

; APPLICANT: Eschenfeldt, William H

; APPLICANT: English, Jami

; APPLICANT: Brinkhaus, Friedhelm L

; TITLE OF INVENTION: Enhanced Carotenoid Accumulation

; TITLE OF INVENTION: in Storage Organs of Genetically

; TITLE OF INVENTION: Engineered Plants

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amoco Corporation, Law Dept

; STREET: 55 Shuman Boulevard, Suite 600

; CITY: Naperville

; STATE: IL

; COUNTRY: USA

; ZIP: 60563-8437

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13937A

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Galloway, Norval B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 7087172447

; TELEFAX: 7087172430

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 177 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-13937A-6

Alignment Scores:

Pred. No.: 2,87e-28 Length: 177

Score: 254.00 Matches: 51

Percent Similarity: 96.6% Conservative: 6

Best Local Similarity: 86.4% Mismatches: 2

Query Match: 83.3% Indels: 0

DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x PCT-US95-13937A-6 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCAGTTCTTTCTCTGAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 60

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AACATGGTGGCGCCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAlaValAlaThrArgSerAsnValThrGlnAla 59

Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATGCAG 177

RESULT 11

US-08-152-483B-8

; Sequence 8, Application US/08152483B

; Patent No. 5529909

; GENERAL INFORMATION:

; APPLICANT: Della-Cioppa, Guy

; APPLICANT: Kumagai, Monto

; TITLE OF INVENTION: TYROSINASE-ACTIVATOR

; TITLE OF INVENTION: PROTEIN FUSION ENZYME

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 2730 Sand Hill Road

; CITY: Menlo Park

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0,

; SOFTWARE: Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/152,483B

; FILING DATE: No. 5529909ember 12, 1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 7/857,602

; FILING DATE: March 30, 1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 923,692

; FILING DATE: July 31, 1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 500,244

; FILING DATE: October 22, 1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 641,617

; FILING DATE: January 16, 1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 737,899

; FILING DATE: July 26, 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Halluin, Albert P.

; REGISTRATION NUMBER: 25,227

; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 854-3660

; TELEFAX: (415) 854-3694

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1442

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: DNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM:

; IMMEDIATE SOURCE:

; CLONE:

; FEATURE:

Alignment Scores:

Pred. No.: 3.18e-26 Length: 1442

Score: 249.00 Matches: 50

Percent Similarity: 96.6% Conservative: 6

Best Local Similarity: 86.2% Mismatches: 2

Query Match: 81.6% Indels: 0

DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x US-08-152-483B-8 (1-1442)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 9 ATGGCTTCTCAGTCTCTTCTCTCGACGAGTTGCCACCCGACGAATGTTGCTCAAGCT 68

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 69 AACATGGTTGCACCTTTCACCTGGCCTTAAGTCAGCTGCCTCATTCCCTGTTTCAAGGAAG 128

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMet 58

Db 129 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATG 182

RESULT 12

US-09-441-340-9

; Sequence 9, Application US/09441340

; Patent No. 6448476

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; TITLE OF INVENTION: Phosphonate Metabolizing Plants

; FILE REFERENCE: 38-21(15303)

; CURRENT APPLICATION NUMBER: US/09/441,340

; CURRENT FILING DATE: 1999-11-16

; EARLIER APPLICATION NUMBER: 60/108,763

; EARLIER FILING DATE: 1998-11-17

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:transit peptide

; OTHER INFORMATION: coding sequence

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(264)

US-09-441-340-9

Alignment Scores:

Pred. No.: 2.58e-20 Length: 264

Score: 201.00 Matches: 41

Percent Similarity: 80.3% Conservative: 8

Best Local Similarity: 67.2% Mismatches: 10

Query Match: 65.9% Indels: 2

DB: 3 Gaps: 1

US-10-718-311-15 (1-62) x US-09-441-340-9 (1-264)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCAGTCTCTTCTCGCTACTATGTT-----GCCTCTCCGGCTCAGGCC 54

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 55 ACTATGGTCGCTCCTTTCAACGGACTTAAGTCCTCGCTGCTTCCAGCCACCCGCAAG 114

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 115 GCTAACACGACATTACTTCCATCAAGCAACGGCGGAAGAGTTAACTGTCATGCAGGTG 174

Qy 61 Trp 61

Db 175 TGG 177

```
RESULT 13
US-09-186-002-5
; Sequence 5, Application US/09186002B
; Patent No. 6489542
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Method for Transforming Plants to Express
; TITLE OF INVENTION: delta-Endotoxins
; FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
; CURRENT APPLICATION NUMBER: US/09/186,002B
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; TYPE: DNA
; LENGTH: 268
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (1)..(267)
; OTHER INFORMATION: coding sequence for PTPI comprising an Arabidopsis
; OTHER INFORMATION: thaliana ssRUBISCO (SSU) chloroplast targeting
; OTHER INFORMATION: sequence and sequences coding for the first 24
; OTHER INFORMATION: amino acids of ssRUBISCO (SSU) protein
US-09-186-002-5

Alignment Scores:
Pred. No.: 2,64e-20 Length: 268
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 3 Gaps: 1

US-10-718-311-15 (1-62) x US-09-186-002-5 (1-268)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 54
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 55 ACTATGGTCGCTCTTCAACGGACTTAAGTCTCGCTTCCGCTTCCAGCCACCCGCAAG 114
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 115 GCTAACACAGACATTACTTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 174
QY 61 Trp 61
Db 175 TGG 177

RESULT 14
US-08-391-339-9
; Sequence 9, Application US/08391339
; Patent No. 5463175
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr.,
; ADDRESS: Monsanto Co. B44F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

US-10-718-311-15 (1-62) x US-08-391-339-9 (1-279)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 12 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 65
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 66 ACTATGGTCGCTCTTCAACGGACTTAAGTCTCGCTTCCGCTTCCAGCCACCCGCAAG 125
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 126 GCTAACACAGACATTACTTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 185
QY 61 Trp 61
Db 186 TGG 188

RESULT 15
US-08-484-274A-9
; Sequence 9, Application US/08484274A
; Patent No. 5776760
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,339
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/156,968
; FILING DATE:
; APPLICATION NUMBER: US/07/717,370
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner, Dennis R., Jr.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10533)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant)
US-08-391-339-9
```

```
Alignment Scores:
Pred. No.: 2,8e-20 Length: 279
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 2 Gaps: 1

US-10-718-311-15 (1-62) x US-08-391-339-9 (1-279)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 12 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 65
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 66 ACTATGGTCGCTCTTCAACGGACTTAAGTCTCGCTTCCGCTTCCAGCCACCCGCAAG 125
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 126 GCTAACACAGACATTACTTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 185
QY 61 Trp 61
Db 186 TGG 188
```

```
RESULT 15
US-08-484-274A-9
; Sequence 9, Application US/08484274A
; Patent No. 5776760
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/484,274A
; APPLICATION NUMBER: US/08/484,274A
; FILING DATE: 07 June 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOST130 38-21(13560)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)789-2679
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant)
US-08-484-274A-9

Alignment Scores:
Pred. No.:      2.8e-20      Length:      279
Score:          201.00      Matches:      41
Percent Similarity: 80.3%    Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match:      65.9%     Indels:      2
DB:               2        Gaps:       1

US-10-718-311-15 (1-62) x US-08-484-274A-9 (1-279)
Qy      1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db      12 ATGGCTTCTCTATGCTCTCTCCGCTACTAGGTT-----GCCTCTCCGGCTCAGGCC 65
Qy      21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db      66 ACTATGGTCGCTCCCTTTCAACGGACTTAAGTCTCCGCTGCCCTCCAGCCACCAGCAAG 125
Qy      41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db      126 GCTAAACAACGACATTACTTCCATCATCAAGCAACGCGGAGAGATTAACTGCATGCAGGTG 185
Qy      61 Trp 61
Db      186 TGG 188
```

Search completed: July 11, 2006, 21:02:57  
Job time : 55.7059 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 19:48:48 ; Search time 1581.32 Seconds  
(without alignments)  
3288.703 Million cell updates/sec

Title: US-10-718-311-15  
Perfect score: 305  
Sequence: 1 MASSVISAANATRSNVTAQ.....LDITSIASNGRVSCMQVWH 62

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/abs/ABSSWEB spool/US10718311/runat\_11072006\_111523\_5213/app\_query.fasta\_1  
-DB=EST -OFMT=fstap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p  
-USER=US10718311@CN 1 1 12067 @runat\_11072006\_111523\_5213 -NCFU=6 -ICPU=3  
-NAEN TIMEOUT=30 -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_hic:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_gss1:\*  
12: gb\_gss2:\*  
13: gb\_gss3:\*  
14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	297	97.4	205	BE462847	BE462847 EST325226
2	297	97.4	235	AW039464	AW039464 EST281745
3	297	97.4	237	AW039041	AW039041 EST281014
4	297	97.4	238	AW041024	AW041024 EST283888

5	297	97.4	245	1	AJ832092	AJ832092 AJ832092
6	297	97.4	250	7	AW442951	AW442951 EST307881
7	297	97.4	260	7	AW037699	AW037699 EST279328
8	297	97.4	260	7	AW094122	AW094122 EST287302
9	297	97.4	264	7	AW037518	AW037518 EST278774
10	297	97.4	264	7	AW037607	AW037607 EST279065
11	297	97.4	267	1	AJ773928	AJ773928 EST255028
12	297	97.4	277	1	AI773502	AI773502 EST254602
13	297	97.4	279	2	BI931188	BI931188 EST51077
14	297	97.4	279	7	BE462486	BE462486 EST34752
15	297	97.4	283	7	AW094091	AW094091 EST287271
16	297	97.4	286	7	BF051382	BF051382 EST436557
17	297	97.4	288	7	AW037811	AW037811 EST279440
18	297	97.4	294	7	AW040243	AW040243 EST282749
19	297	97.4	329	7	AW038470	AW038470 EST280153
20	297	97.4	330	7	AW092013	AW092013 EST285193
21	297	97.4	351	2	BI129648	BI129648 EST475294
22	297	97.4	356	2	BI927635	BI927635 EST547524
23	297	97.4	357	7	AW443393	AW443393 EST308323
24	297	97.4	367	7	AW039845	AW039845 EST282318
25	297	97.4	370	1	AI773548	AI773548 EST284648
26	297	97.4	372	7	BE463004	BE463004 EST325266
27	297	97.4	373	7	AW093805	AW093805 EST286985
28	297	97.4	375	2	BI123527	BI123527 EST469173
29	297	97.4	377	7	AW038793	AW038793 EST280749
30	297	97.4	382	1	AI777051	AI777051 EST252018
31	297	97.4	385	7	AW092154	AW092154 EST285250
32	297	97.4	393	7	AW040522	AW040522 EST283482
33	297	97.4	394	2	BI128649	BI128649 EST474295
34	297	97.4	395	7	AW217063	AW217063 EST295777
35	297	97.4	401	7	AW093581	AW093581 EST286761
36	297	97.4	404	2	BM535307	BM535307 EST588329
37	297	97.4	407	2	BI929701	BI929701 EST549590
38	297	97.4	408	7	AW038880	AW038880 EST280836
39	297	97.4	412	2	BI928909	BI928909 EST548798
40	297	97.4	416	7	AW094716	AW094716 EST287896
41	297	97.4	417	7	AW933080	AW933080 EST358923
42	297	97.4	420	7	AW041064	AW041064 EST283928
43	297	97.4	421	7	AW443412	AW443412 EST308342
44	297	97.4	421	2	BI133589	BI133589 EST466577
45	297	97.4	421	2	BG642684	BG642684 EST510878

#### ALIGNMENTS

RESULT 1  
LOCUS BE462847 205 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST325226 tomato flower buds 0-3 mm, Cornell University  
Lycopersicon esculentum cDNA clone CTOA16011, mRNA sequence.  
ACCESSION BE462847  
VERSION BE462847.1 GI:9508618  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
ORGANISM Lycopersicon esculentum  
REFERENCE 1 (bases 1 to 205)  
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E.,  
Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and  
Tanksley, S.D.  
TITLE Generation of ESTs from tomato flower tissue, 0-3 mm buds  
JOURNAL Unpublished (1999)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
FEATURES Location/Qualifiers  
source 1..205

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOAL6011"
/tissue_type="flower"
/dev_stages="0-3mm buds"
/clone_lib="tomato flower buds 0-3 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tankesley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,096-30 Length: 205  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x BE462847 (1-205)

```

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 67
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
Db 68 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 127
Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 187
Qy 61 Trp 61
Db 188 TGG 190

```

RESULT 2  
 AW039464  
 LOCUS EST281745 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 DEFINITION clone cLET10112, mRNA sequence.

ACCESSION AW039464  
 VERSION AW039464.1 GI:5898218  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 235)  
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,  
 Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,  
 Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and  
 Giovannoni,J.

TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI

Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES  
 source  
 1..235  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"

```

/clone="cLET10112"
/tissue_type="leaf"
/dev_stages="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX, EcoRI
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

```

Alignment Scores:  
 Pred. No.: 2,516-30 Length: 235  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW039464 (1-235)

```

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 14 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 73
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
Db 74 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 133
Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 134 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 193
Qy 61 Trp 61
Db 194 TGG 196

```

RESULT 3  
 AW039041  
 LOCUS EST281014 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 DEFINITION clone cLET10C2, mRNA sequence.

ACCESSION AW039041  
 VERSION AW039041.1 GI:5897795  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 237)  
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,  
 Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,  
 Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and  
 Giovannoni,J.

TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI

Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES  
 source  
 1..237  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET10C2"  
 /tissue\_type="leaf"



/dev\_stage="4-6 week old plants"  
 /lab\_host="Xll-Blue MRP"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2.6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,54e-30 Length: 237  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW039041 (1-237)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 9 ATGGCTTCCTCTGCTCAATTCCTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 68  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 69 AGCATGGTTGCACCTTCACTTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 128  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 129 CAAACCTTGACATCACTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTTG 188  
 |||||  
 QY 61 Trp 61  
 |||||  
 Db 189 TGG 191

## RESULT 4

AW041024  
 LOCUS EST283888 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 DEFINITION clone cLET615, mRNA sequence.

ACCESSION AW041024  
 VERSION AW041024.1 GI:5899778  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,  
 Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W.,  
 Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and  
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue

JOURNAL Unpublished (1999)

COMMENT Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

source

Location/Qualifiers

1. .238  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande Ptor"  
 /db\_xref="taxon:4081"  
 /clone="cLET615"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="Xll-Blue MRP"

/clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2.6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,55e-30 Length: 238  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW041024 (1-238)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 19 ATGGCTTCCTCTGCTCAATTCCTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 78  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 79 AGCATGGTTGCACCTTCACTTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 138  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 139 CAAACCTTGACATCACTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTTG 198  
 |||||  
 QY 61 Trp 61  
 |||||  
 Db 199 TGG 201

## RESULT 5

AJ832092

LOCUS AJ832092

DEFINITION

AJ832092 Lycopersicon esculentum var. cerasiforme fruit 12 dpa  
 Lycopersicon esculentum var. cerasiforme cDNA clone LEEA3\_1, mRNA  
 sequence.

ACCESSION AJ832092

VERSION AJ832092.1

KEYWORDS GI:53705017

SOURCE EST.

Lycopersicon esculentum var. cerasiforme (Solanum lycopersicum var.  
 cerasiforme)

ORGANISM

Lycopersicon esculentum var. cerasiforme  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

AUTHORS Lemaire-Chamley, M., Petit, J., Garcia, V., Just, D., Baldet, P.,  
 Germain, V., Fagard, M., Moussier, M., Chenicet, C. and Rothan, C.  
 Changes in transcriptional profiles are associated with early fruit  
 tissue specialization in tomato

JOURNAL Plant Physiol. 139 (2), 750-769 (2005)

PUBMED 16183847

COMMENT Contact: Garcia V

Physiologie et Biotechnologie Vegetales

INRA

71 avenue E. Bourleaux, Villenave d'Ornon, 33883, FRANCE.

Location/Qualifiers

1. .245  
 /organism="Lycopersicon esculentum var. cerasiforme"  
 /mol\_type="mRNA"  
 /cultivar="West Virginia 106"  
 /db\_xref="taxon:195583"  
 /clone="LEEAA3\_1"  
 /tissue\_type="Fruit"  
 /dev\_stage="12 dpa"  
 /clone\_lib="Lycopersicon esculentum var. cerasiforme fruit  
 12 dpa"

## ORIGIN

Alignment Scores:		2.65e-30	Length:	245
Pred. No.:	297.00	Matches:	61	
Score:	100.0%	Conservative:	0	
Best Local Similarity:	100.0%	Mismatches:	0	
Query Match:	97.4%	Indels:	0	
DB:	1	Gaps:	0	
US-10-718-311-15 (1-62) x AJ832092 (1-245)				
Qy	1	MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla	20	
Db	63	ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGAGCAATGTTACACAAGCT	122	
Qy	21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40	
Db	123	AGCATGGTTGCACCTTTCAGTGGTCTCAAAATCTTCAGGCACCTTTCCTCTGTACAAAGAAG	182	
Qy	41	GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60	
Db	183	CAAAACCTTGACATCATTCTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG	242	
Qy	61	Trp 61		
Db	243	TGG 245		
RESULT 6				
AW442951				
LOCUS	AW442951	250 bp	mRNA	linear EST 18-MAY-2001
DEFINITION	EST307881 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA			
ACCESSION	AW442951			
VERSION	AW442951.1	GI:6985133		
KEYWORDS	EST.			
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.			
REFERENCE	1 (bases 1 to 250)			
AUTHORS	D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Hansen, T.S., Ronning, C.M., Craven, M.B., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D., Giovannoni, J.J. and Martin, G.B.			
TITLE	Generation of ESTs from tomato callus (mixed elicitor)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: CUGI Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence.			
FEATURES	Location/Qualifiers			
source	1..250			
	/organism="Lycopersicon esculentum"			
	/mol_type="mRNA"			
	/cultivar="Rio Grande PtoR"			
	/db_xref="taxon:4081"			
	/clone="cLET42H1"			
	/tissue_type="leaf"			
	/dev_stage="4-6 week old plants"			
	/lab_host="XLI-Blue MRF,"			
	/clone_lib="tomato mixed elicitor, BTI"			
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.87e-30	Length:	260	
Score:	297.00	Matches:	61	
RESULT 7				
AW037699				
LOCUS	AW037699	260 bp	mRNA	linear EST 18-MAY-2001
DEFINITION	EST279328 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA			
ACCESSION	AW037699			
VERSION	AW037699.1	GI:5896453		
KEYWORDS	EST.			
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.			
REFERENCE	1 (bases 1 to 260)			
AUTHORS	D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.			
TITLE	Generation of ESTs from tomato leaf tissue			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: CUGI Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> 5 prime sequence.			
FEATURES	Location/Qualifiers			
source	1..260			
	/organism="Lycopersicon esculentum"			
	/mol_type="mRNA"			
	/cultivar="Rio Grande PtoR"			
	/db_xref="taxon:4081"			
	/clone="cLET4M17"			
	/tissue_type="leaf"			
	/dev_stage="4-6 week old plants"			
	/lab_host="XLI-Blue MRF,"			
	/clone_lib="tomato mixed elicitor, BTI"			
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.87e-30	Length:	260	
Score:	297.00	Matches:	61	

Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW037699 (1-260)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 DB 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 71  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 DB 72 AGCATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 131  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 DB 132 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 191  
 |||||  
 QY 61 TTP 61  
 |||||  
 DB 192 TGG 194

RESULT 8  
 AW094122  
 LOCUS EST287302 tomato mixed elicitor, BTI Lycopersicon esculentum CDNA  
 DEFINITION clone cLET27G20, mRNA sequence.

ACCESSION AW094122  
 VERSION AW094122.1 GI:6059717  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 260)  
 AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,  
 Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Niernan, W.,  
 Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and  
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI

Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source Location/Qualifiers

1..260  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande Ptor"  
 /db\_xref="taxon:4081"  
 /clone="cLET27G20"  
 /issue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

ORIGIN

Alignment Scores:  
 Pred. No.: 2,876-30 Length: 260  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0

Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0  
 US-10-718-311-15 (1-62) x AW094122 (1-260)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 DB 6 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 65  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 DB 66 ACATGGTTGCACCTTCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 125  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 DB 126 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 185  
 |||||  
 QY 61 TTP 61  
 |||||  
 DB 186 TGG 188

RESULT 9  
 AW037518

LOCUS EST278774 tomato mixed elicitor, BTI Lycopersicon esculentum CDNA  
 DEFINITION clone cLET4N17, mRNA sequence.

ACCESSION AW037518  
 VERSION AW037518.1 GI:5896188  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 264)  
 AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,  
 Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Niernan, W.,  
 Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and  
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI

Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source Location/Qualifiers

1..264  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande Ptor"  
 /db\_xref="taxon:4081"  
 /clone="cLET4N17"  
 /issue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

ORIGIN

Alignment Scores:  
 Pred. No.: 2,926-30 Length: 264  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW037518 (1-264)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 12 ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 71  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 72 AGCATGGTGTGCACCTTTCTACTGGTCTCAAAATCTTCAGGCACCTTCCCTGTTTACAAGAAG 131  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 132 CAARACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGTGATGCAAGTG 191  
 |||||  
 QY 61 Trp 61  
 |||||  
 Db 192 TGG 194

RESULT 10  
 AW037607  
 LOCUS  
 DEFINITION EST279065 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 clone cLET4H5, mRNA sequence.  
 ACCESSION AW037607  
 VERSION AW037607.1 GI:5896361  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Lycopersicon esculentum (Solanum lycopersicum)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 264)  
 D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,  
 Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,  
 Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and  
 Giovannoni,J.  
 TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL  
 COMMENT  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 Location/Qualifiers  
 1..264  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET4H5"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,928-30 Length: 264  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0  
 US-10-718-311-15 (1-62) x AW037607 (1-264)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 6 ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 65  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 66 AGCATGGTGTGCACCTTTCTACTGGTCTCAAAATCTTCAGGCACCTTCCCTGTTTACAAGAAG 125  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 126 CAARACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGTGATGCAAGTG 185  
 |||||  
 QY 61 Trp 61  
 |||||  
 Db 186 TGG 188

RESULT 11  
 AW037928  
 LOCUS  
 DEFINITION EST255028 tomato resistant, Cornell Lycopersicon esculentum cDNA  
 clone cLER8L7, mRNA sequence.  
 ACCESSION AW037928  
 VERSION AW037928.1 GI:5271969  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Lycopersicon esculentum (Solanum lycopersicum)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 267)  
 D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,  
 Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,  
 Giovannoni,J.J. and Martin,G.B.  
 TITLE Generation of ESTs from Pseudomonas resistant tomato  
 JOURNAL  
 COMMENT  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 Location/Qualifiers  
 1..267  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="RII-12 (35S::Pto in Rio Grande x Money Maker)"  
 /db\_xref="taxon:4081"  
 /clone="cLER8L7"  
 /tissue\_type="leaf"  
 /dev\_stage="4-week old"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato resistant, Cornell"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLER - Tomato Pseudomonas Resistant EST library.  
 Directionally cloned cDNAs inserted into pBluescript  
 SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,97e-30 Length: 267  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-718-311-15 (1-62) x AW037928 (1-267)  
 QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 23 ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 82  
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 83 AGCATGGTTGCACCTTCTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTGTAACAAGAAG 142  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 DB 143 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 202  
 QY 61 TTP 61  
 DB 203 TGG 205

RESULT 12  
 LOCUS AI773502 277 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST254602 tomato resistant, Cornell Lycopersicon esculentum cDNA  
 clone cLER7A10, mRNA sequence.

ACCESSION AI773502 GI:5271543  
 VERSION EST.  
 KEYWORDS Lycopersicon esculentum (Solanum lycopersicum)  
 SOURCE Lycopersicon esculentum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 277)  
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,  
 Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,  
 Giovannoni,J.J. and Martin,G.B.  
 Generation of ESTs from Pseudomonas resistant tomato  
 Unpublished (1999)

TITLE  
 JOURNAL  
 COMMENT  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: http://www.genome.clemson.edu/orders/index.html  
 5 prime sequence.

FEATURES  
 source  
 Location/Qualifiers  
 1..277  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Ril-12 (35S::Pto in Rio Grande x Money Maker)"  
 /db\_xref="taxon:4081"  
 /clone="CLER7A10"  
 /tissue\_type="leaf"  
 /dev\_stage="4-week old"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato resistant, Cornell"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; CLER - Tomato Pseudomonas Resistant EST Library.  
 Directionally cloned cDNAs inserted into pBlueScript  
 SK(-) at 5' end with EcoRI and 3' end with XhoI site."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,12e-30 Length: 277  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 1 Gaps: 0

US-10-718-311-15 (1-62) x AI773502 (1-277)

QY 1 MetaLaserSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 5 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTTCACACGCGAGCAATTTACACAAGCT 64  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 65 AGCATGGTTGCACCTTCTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTGTAACAAGAAG 124

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 DB 125 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 184  
 QY 61 TTP 61  
 DB 185 TGG 187

## RESULT 13

LOCUS BI931188 279 bp mRNA linear EST 18-OCT-2001  
 DEFINITION EST551077 tomato flower, 8 mm to preanthesis buds Lycopersicon  
 esculentum cDNA clone cFOCI9C6 5' end, mRNA sequence.

ACCESSION BI931188  
 VERSION BI931188.1 GI:16245660  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 279)

AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,  
 Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,  
 Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -

preanthesis

UNPUBLISHED (2001)

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: http://www.genome.clemson.edu/orders/index.html

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3

Location/Qualifiers

1..279

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="CTOC19C6"

/tissue\_type="flower"

/dev\_stage="buds 8mm to preanthesis"

/clone\_lib="tomato flower, 8 mm to preanthesis buds"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:

XhoI; supplier: Cornell University; sequencing: The

Institute for Genomic Research; Flower buds and flowers

were taken from greenhouse plants (4-8 wks old, TA496).

They were immediately frozen in liquid nitrogen and then

size-separated while remaining frozen."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,15e-30 Length: 279  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x BI931188 (1-279)

QY 1 MetaLaserSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 7 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTTCACACGCGAGCAATTTACACAAGCT 66

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

DB 67 AGCATGGTTGCACCTTCTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTGTAACAAGAAG 126

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 127 CAAACCTTGACATCCTCCATTCTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 186

QY 61 Trp 61  
 |||||  
 Db 187 TGG 189

RESULT 14  
 BE462486 279 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST324752 tomato flower buds 0-3 mm, Cornell University  
 DEFINITION Lycopersicon esculentum cDNA clone cTOA1301, mRNA sequence.

ACCESSION BE462486  
 VERSION BE462486.1 GI:9508255  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 279)

REFERENCE  
 AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E.,  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and  
 Tanksley, S.D.  
 Generation of ESTs from tomato flower tissue, 0-3 mm buds  
 Unpublished (1999)  
 CONTACT: CUGI  
 JOURNAL Clemson University Genomics Institute  
 COMMENT 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

FEATURES  
 source  
 1..279  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOA1301"  
 /tissue\_type="flower"  
 /dev\_stage="0-3mm buds"  
 /clone\_lib="tomato flower buds 0-3 mm, Cornell University"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3 15e-30 Length: 279  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x BE462486 (1-279)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 7 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 66

QY 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40  
 |||||  
 Db 67 AGCATGGTGGACCTTCAGTGGTCTCAAAATCTTCAGGCACCTTCCTCTGTGTACAAAGAAG 126

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 127 CAAACCTTGACATCCTCCATTCTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 186

QY 61 Trp 61  
 |||||  
 Db 187 TGG 189

RESULT 15  
 AW094091

DEFINITION EST287271 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 clone cLET27A10, mRNA sequence.

ACCESSION AW094091  
 VERSION AW094091.1 GI:5059686  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 283)

REFERENCE  
 AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,  
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,  
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and  
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

FEATURES  
 source  
 1..283  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET27A10"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRP"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3 21e-30 Length: 283  
 Score: 297.00 Matches: 61  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 97.4% Indels: 0  
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW094091 (1-283)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 12 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 71

QY 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40  
 |||||  
 Db 72 AGCATGGTGGACCTTCACCTTCACCTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 131

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 132 CAAACCTTGACATCCTCCATTCTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 191

QY 61 Trp 61  
 |||||

Db 192 TGG 194

Search completed: July 11, 2006, 23:37:18  
Job time : 1584.32 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 19:35:08 ; Search time 172.484 Seconds  
(without alignments)  
3759.291 Million cell updates/sec

Title: US-10-718-311-15  
Perfect score: 305  
Sequence: 1 MASSVSSAAVATRSNVTA.....LDITSIASNGRVSRCMQVWH 62

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abes/ABSESS pool/US10718311/runat\_11072006\_111518\_5144/app\_query.fasta.1  
-DB=N Geneseg -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bites START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p  
-USER=US10718311 @CN 1.1 1423 @runat\_11072006\_111518\_5144 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

N Geneseg 8:\*  
1: Geneseg1980s:\*  
2: Geneseg1990s:\*  
3: Geneseg2000s:\*  
4: Geneseg2001as:\*  
5: Geneseg2001bs:\*  
6: Geneseg2002as:\*  
7: Geneseg2002bs:\*  
8: Geneseg2003as:\*  
9: Geneseg2003bs:\*  
10: Geneseg2003cs:\*  
11: Geneseg2003ds:\*  
12: Geneseg2004as:\*  
13: Geneseg2004bs:\*  
14: Geneseg2005s:\*  
15: Geneseg2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	305	100.0	684	ABA91838 Chloropla
2	305	100.0	684	Abk47712 DNA seque
3	305	100.0	684	Adc25988 Tomato/Es

4	305	100.0	684	14	ADV91622	Tomato ru
5	269	88.2	204	6	AAI70688	Tobacco R
6	269	88.2	297	10	ADK59826	Plant DNA
7	269	88.2	377	10	ADK56121	Plant DNA
8	269	88.2	489	10	ADK56120	Plant DNA
9	269	88.2	492	10	ADK76164	DNA homol
10	269	88.2	499	6	AAJ35635	Nicotiana
11	269	88.2	504	10	ADK59824	Plant DNA
12	269	88.2	581	10	ADC75089	N bentham
13	269	88.2	608	10	ADC76944	DNA homol
14	269	88.2	614	10	ADC76956	DNA homol
15	269	88.2	615	10	ADC76949	DNA homol
16	269	88.2	632	10	ADC76953	DNA homol
17	269	88.2	684	10	ADC76948	DNA homol
18	269	88.2	718	10	ADK58382	Plant DNA
19	269	88.2	736	10	ADK54321	Plant DNA
20	269	88.2	736	10	ADK57660	Plant DNA
21	269	88.2	841	11	ADM44879	Insect re
22	269	88.2	847	11	ADM45447	Insect re
23	269	88.2	859	11	ADM45065	Insect re
24	269	88.2	924	10	ADC75566	DNA homol
25	269	88.2	958	10	ADC76165	DNA homol
26	269	86.9	619	11	ADM45158	Insect re
27	265	86.9	668	10	ADK59825	Plant DNA
28	265	86.9	668	11	ADM45687	Insect re
29	264	86.6	609	10	ADC76965	DNA homol
30	254	83.3	177	2	AAT27129	Modified
31	254	83.3	177	2	AAT40794	Chloropla
32	254	83.3	177	2	AAT37096	Ribulose
33	254	83.3	177	2	AAT41744	Tobacco t
34	254	83.3	177	2	AAT91545	Phytoene
35	250.5	82.1	683	10	ADK54330	Plant DNA
36	250.5	82.1	737	10	ADK57662	Plant DNA
37	250.5	82.1	737	10	ADK58383	Plant DNA
38	250.5	82.1	789	11	ADM44820	Insect re
39	250.5	82.1	806	11	ADM45686	Insect re
40	249	81.6	177	2	AAQ13721	Phytoene
41	249	81.6	1442	2	AAQ92327	Chloropla
42	244.5	80.2	655	11	ADM45154	Insect re
43	244.5	80.2	738	10	ADK54333	Plant DNA
44	244.5	80.2	753	10	ADK58381	Plant DNA
45	244.5	80.2	754	10	ADK57645	Plant DNA

## ALIGNMENTS

## RESULT 1

ABA91838

ID ABA91838 standard; DNA; 684 BP.

XX ABA91838;

XX 29-AUG-2003 (revised)

DT 15-MAY-2002 (first entry)

XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.

DE Chloroplast transit peptide; tomato; Rubisco; plant;

XX Chloroplast transit peptide; tomato; Rubisco; plant;

KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;

KW enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

OS Chimeric.

XX Key

FH Location/Qualifiers

FT CDS

FT 1..495

FT /\*tag= a

FT /\*product= "chloroplast-targeted CPL fusion"

FT transit\_peptide 1..186

FT /\*tag= b

FT /\*note= "Tomato Rubisco chloroplast targeting sequence"

FT mat\_peptide 187..492

FT /\*tag= c  
 FT /note= "E. coli chorismate pyruvate lyase"  
 XX

PN WO200194607-A2.

XX 13-DEC-2001.

XX 22-MAY-2001; 2001WO-US016661.

XX 02-JUN-2000; 2000US-0209854P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

XX WPI; 2002-226795/28.

DR P-PSDB; AAM50959.

XX Producing para-hydroxy benzoic acid in green plant, comprises expression  
 PT of unique expression cassette containing gene encoding chorismate  
 PT pyruvate lyase operably linked to specific chloroplast targeting  
 PT sequence.

XX Claim 17; Page 57; 60pp; English.

XX The present nucleotide sequence is that of an open reading frame encoding  
 CC chloroplast-targeted chorismate pyruvate lyase (CPL) fusion protein TP-  
 CC CPL (see AAM50959). PCR (see ABA91841-42) was used to generate a DNA  
 CC fragment corresponding to the transit peptide of tomato Rubisco small  
 CC subunit and the first 4 amino acid residues of mature Rubisco. The DNA  
 CC fragment was ligated into pET24a-CPL, which carries the Escherichia coli  
 CC CPL open reading frame (see ABA91837). The TP-CPL construct is an example  
 CC of expression cassettes of the invention that are designed for the high-  
 CC level production of p-hydroxybenzoic acid (PABA) in higher plants.  
 CC Cleavage of TP-CPL in the chloroplast releases a novel polypeptide (see  
 CC AAM50961) that has full enzyme activity, converting chorismate to PABA. A  
 CC claimed plant comprising a CPL expression cassette is selected from  
 CC soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,  
 CC barley, oats, sorghum, rice, Arabidopsis, sugarcane, canola,  
 CC millet, bean, pea, rye, flax or a forage grass. PABA is a monomeric  
 CC component of liquid crystal polymers which have application in the  
 CC automotive, electrical and other industries. (Updated on 29-AUG-2003 to  
 CC standardise OS field)

XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.14e-33 Length: 684  
 Score: 305.00 Matches: 62  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x ABA91838 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 1 ATGGCTTCTCTGTCTATTTCTTACAGCTGTTCACAGCAGCAAGTGTACACAGCT 60  
 Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40  
 Db 61 AGCATGTTGCACCTTTCACCTGGTCTCAAAATCTTCAGCACCTTTCCTGTTACAAAGAAG 120  
 Qy 41 GlnAnLeuAapIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetClnVal 60  
 Db 121 CAAAACCTTGACATCACTTCCATTCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGTG 180

Qy 61 TrpHis 62

Db 181 TGGCAT 186

RESULT 2

ABK47712

ID ABK47712 standard; DNA; 684 BP.

XX AC ABK47712;

XX 18-JUN-2002 (first entry)

XX DNA sequence of ORF for TP-CPL fusion protein.

XX SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester; plant;  
 KW sinapoylglucose:malate sinapoyltransferase; glycosylated aromatic acid;  
 KW malate conjugated aromatic acid; polymer synthesis; ubiC; TP-CPL; gene;  
 KW carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;  
 KW tomato Rubisco small subunit precursor; rbcS2; mutant; ds.

XX Escherichia coli.

OS Lycopersicon esculentum.

OS Synthetic.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 1..684

FT /\*tag= a

FT /product= "TP-CPL fusion protein"

XX WO200204653-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US021283.

XX 07-JUL-2000; 2000US-0216615P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Flint D, Meyer K, Viitanen PV;

XX WPI; 2002-303779/34.

DR P-PSDB; AAU77942.

PT Producing aromatic acid conjugates, involves contacting glycosylated  
 PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the  
 PT presence of sinapoylglucose:malate sinapoyltransferase.

XX Example; Page 70; 72pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana  
 CC gene (SNG1, sinapoylglucose accumulator 1) encoding  
 CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the  
 CC substitution of a glucose moiety on a glycosylated aromatic acid with a  
 CC malate moiety to form a malate conjugated aromatic acid. The enzyme is  
 CC useful for producing malate conjugated aromatic acids, carboxylic acid  
 CC conjugated aromatic acids or aromatic esters. The malate conjugated  
 CC aromatic acids are useful in the synthesis of various polymers. The  
 CC present sequence representing the ORF (open reading frame) for TP-CPL  
 CC (tomato Rubisco small subunit precursor for rbcS2-Escherichia coli  
 CC chorismate pyruvate lyase (CPL)) fusion protein is described in the  
 CC examples of the present invention

XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.14e-33 Length: 684  
 Score: 305.00 Matches: 62  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x ABK47712 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 |||||

Db 1 ATGGCTTCTCTGTCATTTCTTCCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AGCATGGTTGCACCTTTCACCTGCTCTCAAAATCTTCAGCCACCTTTCCTGTTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHis 62

Db 181 TGGCAT 186

RESULT 3

ID ADC25988 standard; DNA; 684 BP.

XX AC ADC25988;

XX 18-DEC-2003 (first entry)

XX Tomato/Escherichia coli chimeric TP-CPL DNA.

XX UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;

KW liquid crystal polymer; LCP; methylparaben; preservative; food;

KW cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;

KW chorismate pyruvate lyase; plant.

XX Chimeric.

OS Lycopersicon esculentum.

OS Escherichia coli.

XX Key Location/Qualifiers

FT CDS 1..684

FT /\*tag= a

FT /product= "Tomato/Escherichia coli chimeric TP-CPL

FT protein"

XX WO2003066836-A2.

XX 14-AUG-2003.

XX 06-FEB-2003; 2003WO-US005863.

XX 07-FEB-2002; 2002US-0355511P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

XX WPI; 2003-767259/72.

XX P-PSDB; ADC25989.

XX New nucleic acid encoding UDP-glucosyltransferase, useful for preparing

PT cells that produce p-hydroxybenzoic acid glucose ester, also the new

PT enzymes.

XX Example 9; SEQ ID NO 41; 161pp; English.

XX The invention relates to a novel isolated nucleic acid that encodes a UDP

CC -glucosyltransferase. The method of the invention may be used to

CC transform microorganisms or green plant cells so that these produce a

CC higher level of high-volume chemicals or materials, such as pHBA (p-

CC hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a

CC monomer for liquid crystal polymers (LCPs) and starting material for

CC methylparaben. Methylparaben is a preservative commonly used in the food

CC and cosmetic industries. The encoded enzymes may be used for in vitro

CC production of these compounds and for identifying similar enzymes by

CC sequence comparison. The current sequence is that of the

CC tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA

CC of the invention.

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,14e-33 Length: 684

Score: 305.00 Matches: 62

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC25988 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCTGTCATTTCTTCCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AGCATGGTTGCACCTTTCACCTGCTCTCAAAATCTTCAGCCACCTTTCCTGTTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHis 62

Db 181 TGGCAT 186

RESULT 4

ID ADV91622 standard; DNA; 684 BP.

XX AC ADV91622;

XX 10-MAR-2005 (first entry)

XX Tomato rubisco small subunit precursor TP-E. coli CPL chimeric DNA.

XX Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;

KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;

KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;

KW genetically engineered microorganism; antioxidant; antimicrobial agent;

KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;

KW neoplasm; pHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid;

KW rubisco small subunit precursor; transit peptide;

KW chloroplast transit peptide; TP; chloroplast; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 1..684

FT /\*tag= a

FT /product= "Tomato rubisco small subunit precursor TP-E.

FT coli CPL fusion protein"

XX US2004261147-A1.

XX 23-DEC-2004.

XX 16-JUN-2003; 2003US-00462162.

XX 16-JUN-2003; 2003US-00462162.

XX (MEYE/) MEYER K.

XX (VIIT/) VIITANEN P V.

XX (FLIN/) FLINT D.

XX Meyer K, Viitanen PV, Flint D;

XX WPI; 2005-057232/06.

XX P-PSDB; ADV91623.

XX Producing hydroquinone glucoside in a green plant comprises growing a  
 PT green plant having nucleic acid fragments and chorismate pyruvate lyase  
 PT expression cassette.  
 XX  
 XX  
 PS Example 4; SEQ ID NO 33; 70pp; English.  
 XX  
 XX The present invention relates to methods and materials to produce  
 CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
 CC in genetically modified green plants and microorganisms. The method  
 CC relies upon transgenic plants or genetically modified microorganisms that  
 CC produce increased levels of the initial substrate para-hydroxybenzoic  
 CC acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce  
 CC arbutin. Plants and microbes can be genetically engineered to produce  
 CC high levels of pHBA either by functional expression of the bacterial  
 CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-  
 CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
 CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
 CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
 CC present sequence is tomato rubisco small subunit precursor (rbcS2)  
 CC transit peptide (chloroplast transit peptide; TP)-Escherichia coli CPL  
 CC chimeric DNA. This sequence is present in the E. coli expression vector  
 CC construct pET24a-TP-CPL.

XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5,148-33 Length: 684  
 Score: 305.00 Matches: 62  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 14 Gaps: 0

US-10-718-311-15 (1-62) x ADV91622 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 1 ATGGCTTCTCTGTCATTCTTTCAGCAGCTGTTGCCACGCGAGCAATGTTACACAAGCT 60  
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyss 40  
 Db 61 AGCATGGGTGGACCTTTCAGTGGTCTCAAAATCTTCAGGCACCTTCCCTGTTACAAAGAAG 120  
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 121 CAAACCTTGACATCACTTCCATTCTAGCATGTTGGAGAGTTAGTGTGATGCGAGTG 180  
 Qy 61 TrpHis 62  
 Db 181 TGGCAT 186

RESULT 5

AAI70688

ID AAI70688 standard; DNA; 204 BP.

XX

AC AAI70688;

XX

DT 04-FEB-2002 (first entry)

XX

DE Tobacco Rubisco transit peptide small subunit DNA.

XX

KW Tobacco; Rubisco; ribulosebiphosphate carboxylase; transit peptide;

XX

KW acetyl-CoA carboxylase; transgenic plant; oilseed; vegetable oil;

XX

XX Brassica; soybean; ds.

OS Nicotiana tabacum.

XX

Key Location/Qualifiers

FT transit\_peptide 1..171

FT /\*tag= a

FT mat\_peptide 172..204

FT /\*tag= b

FT /note= "5' end of mature protein coding sequence"  
 XX  
 FN WO2000181604-A1.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US012928.  
 XX  
 PR 20-APR-2000; 2000US-0198794P.  
 XX  
 PA (CRGI ) CARGILL INC.  
 XX  
 PI Shorresh BS, Debonte LR;  
 XX  
 DR WPI; 2002-041417/05.  
 DR P-PSDB; AAM50331.  
 XX  
 PT Using Nucleic acids encoding alfalfa cytosolic Acetyl-Coenzyme A  
 PT Carboxylase (ACCase; EC 6.4.1.2) to increase the oil content of Soybean  
 PT and Brassica seeds.  
 XX  
 PS Disclosure; Fig 1; 62pp; English.

XX The present sequence is that of DNA encoding the tobacco small subunit  
 CC Rubisco transit peptide and the 5' portion of the mature small subunit  
 CC Rubisco protein (see AAM50331). Nucleic acid constructs of the invention  
 CC include a nucleic acid encoding a cytosolic acetyl-CoA carboxylase (EC-  
 CC 6.4.1.2) and a promoter, and may also include a nucleic acid encoding the  
 CC tobacco small subunit Rubisco transit peptide. Such constructs can be  
 CC introduced into soybean and Brassica plants (especially Brassica napus  
 CC (oilseed rape), Brassica rapa (turnip), Brassica juncea, Brassica  
 CC carinata, Brassica nigra (black mustard) and Brassica oleracea  
 CC (cauliflower, sprout, cabbage, broccoli)) to increase the oil content of  
 CC their seeds from about 5% to about 25% greater on a dry weight basis  
 CC (claimed)

SQ Sequence 204 BP; 50 A; 60 C; 43 G; 51 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,11e-28 Length: 204  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x AAI70688 (1-204)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 1 ATGGCTTCTCTGTCATTCTTTCAGCAGCTGTTGCCACGCGAGCAATGTTGCTCAAGCT 60  
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyss 40  
 Db 61 AACATGGTGGACCTTTCAGTGGCTTAAATGTCAGCTGCTCATCCCTGTTTCAGGAAG 120  
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACGCGAGAGTGCATGTCAGGTG 180  
 Qy 61 Trp 61  
 Db 181 TGG 183

RESULT 6

ADK59826

ID ADK59826 standard; DNA; 297 BP.

XX

AC ADK59826;

XX

DT 06-MAY-2004 (first entry)

XX

DE Plant DNA sequence which confers altered metabolic characteristic #7209.

XX altered metabolic characteristic; plant; acid metabolism;  
 KW alcohol metabolism; fatty acid metabolism;  
 KW branched fatty acid metabolism; alkaloïd metabolism;  
 KW amino acid metabolism; ester metabolism; glyceride metabolism;  
 KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;  
 KW terpene metabolism; isoprenoid metabolism; alkene metabolism;  
 KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;  
 KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.  
 XX Unidentified.  
 OS  
 XX  
 XX WO2003020936-A1.  
 XX  
 XX 13-MAR-2003.  
 XX  
 XX 30-AUG-2002; 2002WO-US027884.  
 XX  
 XX 31-AUG-2001; 2001US-0316471P.  
 XX  
 XX (DOWC ) DOW CHEM CO.  
 XX (DOWC ) DOW AGROSCIENCES LLC.  
 XX  
 XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;  
 PI Oriado JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;  
 XX  
 XX WPI; 2003-313091/30.  
 XX  
 XX Novel genes that confer altered metabolic characteristics in Nicotiana  
 PT benthamiana plants, useful for altering the levels of metabolites e.g.  
 PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.  
 XX  
 XX Claim 1; SEQ ID NO 7209; 2576pp; English.  
 XX  
 XX The invention comprises DNA sequences which confer an altered metabolic  
 CC characteristic when they are expressed in a plant. The DNA sequences of  
 CC the invention are useful for producing plants with an altered metabolic  
 CC characteristic, such as: altered acid metabolism, alcohol metabolism,  
 CC fatty acid metabolism, branched fatty acid metabolism, alkaloïd or other  
 CC base metabolism, altered amino acid metabolism, altered ester metabolism,  
 CC altered glyceride metabolism, altered phenolic metabolism, altered  
 CC carbohydrate metabolism, altered sterol, oxygenated terpene, or  
 CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon  
 CC metabolism, ketone or quinone metabolism. The DNA sequences of the  
 CC invention may be used to provide disease resistance in a plant and gene  
 CC shuffling or sexual PCR procedures. The present nucleic acid represents a  
 CC DNA sequence of the invention.  
 XX  
 XX Sequence 297 BP; 83 A; 70 C; 70 G; 74 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.88e-28 Length: 297  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 10 Indels: 0  
 DB: 8 Gaps: 0  
 US-10-718-311-15 (1-62) x ADK59826 (1-297)  
 QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 Db 97 ATGGCTTCTCAGTCTTCTTCTCAGCAGCAGTTCGCCACCGCAGCATGTGCTCAAGCT 156  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 Db 157 AACATGGTTCACCTTTCACAGCTCTTAAGTCTGCTGCCTCATTCCTCTTTCACGAAG 216  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 Db 217 CAAACCTTGACATCACTTCCATTGCGCAGCAACGGGGAAGAGTGAATGCATGCAGGTG 276  
 QY 61 Trp 61

Db 277 TGG 279  
 RESULT 7  
 ADK56121  
 ID ADK56121 standard; DNA; 377 BP.  
 XX  
 AC ADK56121;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 XX Plant DNA sequence which confers altered metabolic characteristic #3504.  
 XX altered metabolic characteristic; plant; acid metabolism;  
 KW alcohol metabolism; fatty acid metabolism;  
 KW branched fatty acid metabolism; alkaloïd metabolism;  
 KW amino acid metabolism; ester metabolism; glyceride metabolism;  
 KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;  
 KW terpene metabolism; isoprenoid metabolism; alkene metabolism;  
 KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;  
 KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.  
 XX Unidentified.  
 OS  
 XX  
 XX WO2003020936-A1.  
 XX  
 XX 13-MAR-2003.  
 XX  
 XX 30-AUG-2002; 2002WO-US027884.  
 XX  
 XX 31-AUG-2001; 2001US-0316471P.  
 XX  
 XX (DOWC ) DOW CHEM CO.  
 XX (DOWC ) DOW AGROSCIENCES LLC.  
 XX  
 XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;  
 PI Oriado JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;  
 XX  
 XX WPI; 2003-313091/30.  
 XX  
 XX Novel genes that confer altered metabolic characteristics in Nicotiana  
 PT benthamiana plants, useful for altering the levels of metabolites e.g.  
 PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.  
 XX  
 XX Claim 1; SEQ ID NO 3504; 2576pp; English.  
 XX  
 XX The invention comprises DNA sequences which confer an altered metabolic  
 CC characteristic when they are expressed in a plant. The DNA sequences of  
 CC the invention are useful for producing plants with an altered metabolic  
 CC characteristic, such as: altered acid metabolism, alcohol metabolism,  
 CC fatty acid metabolism, branched fatty acid metabolism, alkaloïd or other  
 CC base metabolism, altered amino acid metabolism, altered ester metabolism,  
 CC altered glyceride metabolism, altered phenolic metabolism, altered  
 CC carbohydrate metabolism, altered sterol, oxygenated terpene, or  
 CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon  
 CC metabolism, ketone or quinone metabolism. The DNA sequences of the  
 CC invention may be used to provide disease resistance in a plant and gene  
 CC shuffling or sexual PCR procedures. The present nucleic acid represents a  
 CC DNA sequence of the invention.  
 XX  
 XX Sequence 377 BP; 103 A; 89 C; 86 G; 99 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2.62e-28 Length: 377  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 10 Indels: 0  
 DB: 18 Gaps: 0  
 US-10-718-311-15 (1-62) x ADK56121 (1-377)  
 QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

```
Db 68 ATGGCTTCCTCAGTTCTTCCTCAGCAGCAGTTGCACCCGCGAGCAATGTTGCTCAAGCT 127
Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaAlaThrPheProValThrIysLys 40
Db 128 AACATGGTGGTGCACCTTTCACAGGCTTAAAGTCTGCTGCCTCATTCCTCTGTTTCAAGAAAG 187
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 188 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGAGTGCATGCAATGCATGAGGTG 247
Qy 61 Trp 61
Db 248 TGG 250

RESULT 8
ADK56120
ID ADK56120 standard; DNA; 489 BP.
XX
AC ADK56120;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant DNA sequence which confers altered metabolic characteristic #3503.
XX
KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkaloid metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; steroid metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
OS Unidentified.
XX
PN WO2003020936-A1.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
PR 31-AUG-2001; 2001US-0316471P.
XX
PA (DOWC ) DOW CHEM CO.
XX
PA (DOWC ) DOW AGROSCIENCES LLC.
XX
PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ,
PI Orledo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA,
XX
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
XX benthamiana plants, useful for altering the levels of metabolites e.g
XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
PS Claim 1; SEQ ID NO 3503; 2576pp; English.
XX
XX The invention comprises DNA sequences which confer an altered metabolic
XX characteristic when they are expressed in a plant. The DNA sequences of
XX the invention are useful for producing plants with an altered metabolic
XX characteristic, such as: altered acid metabolism, alcohol metabolism,
XX fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
XX base metabolism, altered amino acid metabolism, altered ester metabolism,
XX altered glyceride metabolism, altered phenolic metabolism, altered
XX carbohydrate metabolism, altered steroid, oxygenated terpene, or
XX isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
XX metabolism, ketone or quinone metabolism. The DNA sequences of the
XX invention may be used to provide disease resistance in a plant and gene
XX shuffling or sexual PCR procedures. The present nucleic acid represents a
XX DNA sequence of the invention.
XX
XX Sequence 489 BP; 136 A; 113 C; 114 G; 126 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 3,77e-28 Length: 489
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADK56120 (1-489)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 79 ATGGCTTCCTCAGTTCTTCCTCAGCAGCAGTTGCACCCGCGAGCAATGTTGCTCAAGCT 138
Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaAlaThrPheProValThrIysLys 40
Db 139 AACATGGTGGTGCACCTTTCACAGGCTTAAAGTCTGCTGCCTCATTCCTCTGTTTCAAGAAAG 198
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 199 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGCAATGCATGAGGTG 258
Qy 61 Trp 61
Db 259 TGG 261

RESULT 9
ADK76164
ID ADC76164 standard; DNA; 492 BP.
XX
AC ADC76164;
XX
DT 01-JAN-2004 (first entry)
XX
DE DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1433.
XX
KW rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW phytopathogen; gene shuffling; ds.
XX
OS Unidentified.
XX
PN WO2003020905-A2.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027883.
XX
PR 31-AUG-2001; 2001US-0316392P.
XX
XX (DOWC ) DOW CHEM CO.
XX
XX Shukla V, Butler H, Larrinua I, Reddy AS;
XX WPI; 2003-290185/28.
XX
XX Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
XX sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
XX rhoeas, useful for conferring disease resistance in plants.
XX
XX Claim 1; SEQ ID NO 1433; 617pp; English.
XX
XX The invention relates to a novel isolated nucleic acid derived from
XX Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
XX (yeast), Trichoderma harzianum (Hypocrea lixi) and Papaver rhoeas
XX (poppy) and a sequence that hybridises to them under conditions of low
XX stringency, where expression of the nucleic acid in a plant results in a
XX disease resistance phenotype. The polynucleotides of the invention
XX demonstrate anti-fungal activity and may be useful in conferring disease
XX resistance in a plant against phytopathogen such as Aspergillus flavus,
XX Gibberella fujikuroi and Gibberella zeae. Furthermore, the
XX polynucleotides may be useful to retrieve unknown sequences and in gene
XX shuffling or sexual PCR procedures. The current sequence is that of the
```

CC DNA of the invention which is homologous to that of the phytopathogen  
CC resistance-related contig cDNAs.  
SQ Sequence 492 BP; 140 A; 115 C; 111 G; 126 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.8e-28 Length: 492  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 10 Gaps: 0  
US-10-718-311-15 (1-62) x ADC76164 (1-492)  
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 66 ATGGCTTCTCAGTCTTCTCCTCAGCAGCAGTGTGCCACCGCAGCAATGTTGCTCAAGCT 125  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 126 AACATGGTGGCACCCTTTCACCTGCTTAAGTCTGCTGCCTCATTCCTCTGTTTCAAGGAAG 185  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
Db 186 CAAACCTTGACATCACTTCCATTGCCAGTGCAGACGGCGGAAGAGTGCAATGCATGCAGGTG 245  
QY 61 TTP 61  
Db 246 TGG 248  
RESULT 10  
AAD35635  
ID AAD35635 standard; cDNA; 499 BP.  
XX  
AC AAD35635;  
XX  
DT 26-JUL-2002 (first entry)  
XX  
DE Nicotiana benthamiana rubisco partial cDNA sequence.  
XX  
KW Nucleic acid vector; tobacco rattle virus; virus induced gene silencing;  
KW TRV; VIGS; ribulose-1,5-bisphosphate carboxylase oxygenase; rubisco;  
KW transgenic plant; enzyme; ss.  
XX  
OS Nicotiana benthamiana.  
XX  
PN US6369296-B1.  
XX  
PD 09-APR-2002.  
XX  
PF 01-FEB-2000; 2000US-00495797.  
XX  
PR 01-FEB-2000; 2000US-00495797.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Ratcliff FG, Martin-Hernandez AM, Baulcombe DC;  
XX  
DR WPI; 2002-337969/37.  
XX  
PT Novel nucleic acid vector, useful for producing transgenic plants,  
PT comprises a plant active promoter linked to a recombinant tobacco rattle  
PT virus cDNA.  
XX  
PS Example 2; Col 37-38; 32pp; English.  
XX  
CC The invention relates to a nucleic acid vector. The vector comprises a  
CC plant active promoter operably linked to a recombinant tobacco rattle  
CC virus (TRV) cDNA which includes cis acting elements of TRV RNA permitting  
CC in the presence of replicase, replication of the cDNA, a viral subgenomic  
CC promoter operably linked to a sequence encoding a TRV coat protein, and a  
CC heterologous sequence, and border sequences for transfer into a plant

CC genome. The vector is useful for producing genetically engineered plant  
CC cells or transgenic plants. The nucleic acids are used to initiate  
CC virus induced gene silencing (VIGS). The present sequence is Nicotiana  
CC benthamiana ribulose-1,5-bisphosphate carboxylase oxygenase (rubisco)  
CC partial cDNA related to the invention  
XX  
SQ Sequence 499 BP; 132 A; 120 C; 123 G; 124 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.88e-28 Length: 499  
Score: 269.00 Matches: 53  
Percent Similarity: 96.7% Conservative: 6  
Best Local Similarity: 86.9% Mismatches: 2  
Query Match: 88.2% Indels: 0  
DB: 6 Gaps: 0  
US-10-718-311-15 (1-62) x AAD35635 (1-499)  
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 8 ATGGCTTCTCAGTCTTCTCCTCAGCAGTGTGCCACCGCAGCAATGTTGCTCAAGCT 67  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 68 ACATGGTGGCACCCTTTCACCTGCTTAAGTCACTGCTCATTCCTCTGTTTCAAGGAAG 127  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
Db 128 CAAACCTTGACATCACTTCCATTGCCAGTGCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 187  
QY 61 TTP 61  
Db 188 TGG 190  
RESULT 11  
ADK59824  
ID ADK59824 standard; DNA; 504 BP.  
XX  
AC ADK59824;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Plant DNA sequence which confers altered metabolic characteristic #7207.  
XX  
KW altered metabolic characteristic; plant; acid metabolism;  
KW alcohol metabolism; fatty acid metabolism;  
KW branched fatty acid metabolism; alkaloid metabolism;  
KW amino acid metabolism; ester metabolism; glyceride metabolism;  
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;  
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;  
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;  
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.  
XX  
OS Unidentified.  
XX  
PN WO2003020936-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 30-AUG-2002; 2002WO-US027884.  
XX  
PR 31-AUG-2001; 2001US-031647P.  
XX  
PA (DOWC) DOW CHEM CO.  
PA (DOWC) DOW AGROSCIENCES LLC.  
XX  
PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;  
PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;  
XX WPI; 2003-313091/30.  
XX  
PT Novel genes that confer altered metabolic characteristics in Nicotiana  
PT benthamiana plants, useful for altering the levels of metabolites e.g.

PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.  
 XX  
 PS Claim 1; SEQ ID NO 7207; 2576pp; English.  
 XX  
 CC The invention comprises DNA sequences which confer an altered metabolic  
 CC characteristic when they are expressed in a plant. The DNA sequences of  
 CC the invention are useful for producing plants with an altered metabolic  
 CC characteristic, such as: altered acid metabolism, alcohol metabolism,  
 CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other  
 CC base metabolism, altered amino acid metabolism, altered ester metabolism,  
 CC altered glyceride metabolism, altered phenolic metabolism, altered  
 CC carbohydrate metabolism, altered sterol, oxygenated terpene, or  
 CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon  
 CC metabolism, ketone or quinone metabolism. The DNA sequences of the  
 CC invention may be used to provide disease resistance in a plant and gene  
 CC shuffling or sexual PCR procedures. The present nucleic acid represents a  
 CC DNA sequence of the invention.  
 XX  
 SQ Sequence 504 BP; 139 A; 117 C; 118 G; 130 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.93e-28 Length: 504  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADK59824 (1-504)

QY 1 MetAlaSerSerValIleSerSerAlaLaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 78 ATGGCTTCTCAGTCTTCTCTCAGCAGCAGTTCGCCACCCGAGCAATGTTGCTCAAGCT 137  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 138 AACATGGTTGCACCTTTCACAGGTCTTAAGTCTGCTGCTCATTCCCTGTTTCAAGAAAG 197  
 QY 41 GlnAnLeuAspIleThrSerIleAlaSerAnGlyGlyArgValSerCysMetGlnVal 60  
 DB 198 CAAACCTTGACATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGCATGCAGGTG 257  
 QY 61 Trp 61  
 DB 258 TGG 260

RESULT 12  
 ADC75089  
 ID ADC75089 standard; cDNA; 581 BP.  
 XX  
 AC ADC75089;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE N benthamiana phytopathogen resistance-related contig cDNA - SEQ ID 13.  
 XX  
 KW rice; yeast; poppy; plant; disease resistance; anti-fungal;  
 KW phytopathogen; gene shuffling; ss.  
 XX  
 OS Nicotiana benthamiana.  
 XX  
 PN WO2003020905-A2.  
 XX  
 PD 13-MAR-2003.  
 XX  
 PF 30-AUG-2002; 2002WO-US027883.  
 XX  
 PR 31-AUG-2001; 2001US-0316392P.  
 XX  
 PA (DOWC) DOW CHEM CO.  
 XX  
 PI Shukla V, Butler H, Larrinua I, Reddy AS;  
 XX

DR WPI; 2003-290185/28.  
 XX  
 PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae  
 PT sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver  
 PT rhoeas, useful for conferring disease resistance in plants.  
 XX  
 PS Claim 1; SEQ ID NO 13; 617pp; English.  
 XX  
 CC The invention relates to a novel isolated nucleic acid derived from  
 CC Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae  
 CC (yeast), Trichoderma harzianum (Hypocrea lixi) and Papaver rhoeas  
 CC (poppy) and a sequence that hybridises to them under conditions of low  
 CC stringency, where expression of the nucleic acid in a plant results in a  
 CC disease resistance phenotype. The polynucleotides of the invention  
 CC demonstrate anti-fungal activity and may be useful in conferring disease  
 CC resistance in a plant against phytopathogen such as Aspergillus flavus,  
 CC Gibberella fujikuroi and Gibberella zeae. Furthermore, the  
 CC polynucleotides may be useful to retrieve unknown sequences and in gene  
 CC shuffling or sexual PCR procedures. The current sequence is that of the  
 CC phytopathogen resistance-related contig cDNA of the invention.  
 XX  
 SQ Sequence 581 BP; 162 A; 132 C; 143 G; 144 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.8e-28 Length: 581  
 Score: 269.00 Matches: 53  
 Percent Similarity: 96.7% Conservative: 6  
 Best Local Similarity: 86.9% Mismatches: 2  
 Query Match: 88.2% Indels: 0  
 DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC75089 (1-581)

QY 1 MetAlaSerSerValIleSerSerAlaLaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 63 ATGGCTTCTCAGTCTTCTCTCAGCAGCAGTTCGCCACCCGAGCAATGTTGCTCAAGCT 122  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 123 AACATGGTTGCACCTTTCACAGGTCTTAAGTCTGCTGCTCATTCCCTGTTTCAAGAAAG 182  
 QY 41 GlnAnLeuAspIleThrSerIleAlaSerAnGlyGlyArgValSerCysMetGlnVal 60  
 DB 183 CAAACCTTGACATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGCATGCAGGTG 242  
 QY 61 Trp 61  
 DB 243 TGG 245

RESULT 13  
 ADC76944  
 ID ADC76944 standard; DNA; 608 BP.  
 XX  
 AC ADC76944;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1213.  
 XX  
 KW rice; yeast; poppy; plant; disease resistance; anti-fungal;  
 KW phytopathogen; gene shuffling; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003020905-A2.  
 XX  
 PD 13-MAR-2003.  
 XX  
 PF 30-AUG-2002; 2002WO-US027883.  
 XX  
 PR 31-AUG-2001; 2001US-0316392P.  
 XX  
 PA (DOWC) DOW CHEM CO.



```

XX  Shukla V, Butler H, Larrinua I, Reddy AS;
PI  WPI; 2003-290185/28.
XX
XX  Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
PT  sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT  rhoas, useful for conferring disease resistance in plants.
XX
XX  Claim 1; SEQ ID NO 1213; 617pp; English.
XX
XX  The invention relates to a novel isolated nucleic acid derived from
CC  Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC  (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoas
CC  (poppy) and a sequence that hybridises to them under conditions of low
CC  stringency, where expression of the nucleic acid in a plant results in a
CC  disease resistance phenotype. The polynucleotides of the invention
CC  demonstrate anti-fungal activity and may be useful in conferring disease
CC  resistance in a plant against phytopathogen such as Aspergillus flavus,
CC  Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC  polynucleotides may be useful to retrieve unknown sequences and in gene
CC  shuffling or sexual PCR procedures. The current sequence is that of the
CC  DNA of the invention which is homologous to that of the phytopathogen
CC  resistance-related contig cDNAs.
XX
SQ  Sequence 608 BP; 164 A; 144 C; 142 G; 158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,11e-28 Length: 608
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76944 (1-608)

QY  1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB  38 ATGGCTTCCTCAGTTCTTTCTCAGCAGCAGTTCGCCACCGCAGCAATGTGCTCAAGCT 97
QY  21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB  98 AACATGGTGCACCTTTACAGGTCCTTAAGTCGTGCTCCTCATCTCCCTGTTTCAAGAAAG 157
QY  41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB  158 CAAACCTTGACATCACTTCATTCAGCAGCAGTTCGCCACCGCGGAAGAGTGCAATGCATGCAGGTG 217
QY  61 Trp 61
DB  218 TGG 220

RESULT 14
ADC76956
ID  ADC76956 standard; DNA; 614 BP.
XX
XX  ADC76956;
AC
AC  01-JAN-2004 (first entry)
DT
DE  DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1225.
XX
XX  rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW  phytopathogen; gene shuffling; ds.
XX
XX  Unidentified.
OS
XX  WO2003020905-A2.
XX
XX  13-MAR-2003.
XX
XX  30-AUG-2002; 2002WO-US027883.
PF

```

```

XX  31-AUG-2001; 2001US-0316392P.
XX
XX  (IDWC ) DOW CHEM CO.
XX
XX  Shukla V, Butler H, Larrinua I, Reddy AS;
PI  WPI; 2003-290185/28.
XX
XX  Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
PT  sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT  rhoas, useful for conferring disease resistance in plants.
XX
XX  Claim 1; SEQ ID NO 1225; 617pp; English.
XX
XX  The invention relates to a novel isolated nucleic acid derived from
CC  Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC  (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoas
CC  (poppy) and a sequence that hybridises to them under conditions of low
CC  stringency, where expression of the nucleic acid in a plant results in a
CC  disease resistance phenotype. The polynucleotides of the invention
CC  demonstrate anti-fungal activity and may be useful in conferring disease
CC  resistance in a plant against phytopathogen such as Aspergillus flavus,
CC  Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC  polynucleotides may be useful to retrieve unknown sequences and in gene
CC  shuffling or sexual PCR procedures. The current sequence is that of the
CC  DNA of the invention which is homologous to that of the phytopathogen
CC  resistance-related contig cDNAs.
XX
SQ  Sequence 614 BP; 167 A; 145 C; 149 G; 153 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,18e-28 Length: 614
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76956 (1-614)

QY  1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB  46 ATGGCTTCCTCAGTTCTTTCTCAGCAGCAGTTCGCCACCGCAGCAATGTGCTCAAGCT 105
QY  21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB  106 AACATGGTTCACCTTTACATTCGTCTTAAGTCAGTCGCTCCTGTTTCAAGGAAG 165
QY  41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB  166 CAAACCTTGACATCACTTCATTCAGCAGCAGTTCGCCACCGCGGAAGAGTGCAATGCATGCAGGTG 225
QY  61 Trp 61
DB  226 TGG 228

RESULT 15
ADC76949
ID  ADC76949 standard; DNA; 615 BP.
XX
XX  ADC76949;
AC
AC  01-JAN-2004 (first entry)
DT
DE  DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1218.
XX
XX  rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW  phytopathogen; gene shuffling; ds.
XX
XX  Unidentified.
OS
XX  WO2003020905-A2.
PF

```

```

XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027893.
XX
XX 31-AUG-2001; 2001US-0316392P.
XX
XX (DOWC ) DOW CHEM CO.
XX
XX Shukla V, Butler H, Larrinua I, Reddy AS;
XX
XX WPI; 2003-290185/28.
XX
XX Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
XX sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
XX rhoeas, useful for conferring disease resistance in plants.
XX
XX Claim 1; SEQ ID NO 1218; 617pp; English.
XX
XX The invention relates to a novel isolated nucleic acid derived from
XX Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
XX (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoeas
XX (poppy) and a sequence that hybridises to them under conditions of low
XX stringency, where expression of the nucleic acid in a plant results in a
XX disease resistance phenotype. The polynucleotides of the invention
XX demonstrate anti-fungal activity and may be useful in conferring disease
XX resistance in a plant against phytopathogen such as Aspergillus flavus,
XX Gibberella fujikuroi and Gibberella zeae. Furthermore, the
XX polynucleotides may be useful to retrieve unknown sequences and in gene
XX shuffling or sexual PCR procedures. The current sequence is that of the
XX DNA of the invention which is homologous to that of the phytopathogen
XX resistance-related contig cDNAs.
XX
XX SQ Sequence 615 BP; 168 A; 144 C; 146 G; 156 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:          5.2e-28          Length:          615
Score:             269.00           Matches:         53
Percent Similarity: 96.7%           Conservative:    6
Best Local Similarity: 86.9%         Mismatches:     2
Query Match:       88.2%             Indels:         0
DB:                10               Gaps:           0

US-10-718-311-15 (1-62) x ADC76949 (1-615)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 47 ATGGCTTCCTCAGTCTTTCTCCAGCAGCAGTTCGCCGCCGAGCAATGTGCTCAAGCT 106
QY 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40
Db 107 AACATGGTTGCACCTTTCACAGCTCTTAAGTCTGCTGCTCATTCCTGTTTCAAGAAAG 166
QY 41 GlnAsnLeuAepIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 167 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTCAATGCATGCAGGTG 226
QY 61 Trp 61
Db 227 TGG 229

Search completed: July 11, 2006, 21:50:41
Job time : 173.484 secs

```

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM protein - nucleic search, using frame\_plus\_p2n model  
 Run on: July 11, 2006, 20:58:40 ; Search time 135.1 Seconds  
 (without alignments)  
 3276.788 Million cell updates/sec

Title: US-10-718-311-8  
 Sequence: 1 MASSVISAAVATRSNVTOA.....RLSGPRLLTFLPASPPLY 227

Scoring table: BLOSUM62  
 Xgapop 10.0, Xgapext 0.5  
 Ygapop 10.0, Ygapext 0.5  
 Fgapop 6.0, Fgapext 7.0  
 Delop 6.0, Delext 7.0

Searched: 844069 seqs, 650066433 residues  
 Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Command line parameters:  
 -MODEL=frame+ p2n.model -DEV=xlp  
 -Q=/abes/ABSWEB/spool/US10718311/runat\_11072006\_111535\_5434/app\_query.fasta.1  
 -DB=Published Applications NA New -QWMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
 -MAXLEN=2000000000 -HOST=abs06p  
 -USER=US10718311 @CGN 1.1.251 @runat\_11072006\_111535\_5434 -NCPU=6 -ICPU=3  
 -NO MAP -NEG SCORES=0 -WAIT -DSPELQ=100 -LONGLOG -DEV TIMEOUT=120  
 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:  
 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
 9: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	281.5	24.3	781	US-11-247-587-124	Sequence 124, App
2	276.5	23.9	782	US-11-247-587-125	Sequence 125, App
3	253.5	21.9	698	US-11-247-587-126	Sequence 126, App
4	247.5	21.4	683	US-11-247-587-127	Sequence 127, App
5	246.5	21.3	700	US-11-247-587-127	Sequence 127, App
6	238.5	20.6	583	US-11-218-305-17631	Sequence 17631, A
7	231.5	20.0	727	US-11-247-587-129	Sequence 129, App
8	228.5	19.7	743	US-10-953-349-8894	Sequence 8894, App

9	227.5	19.6	558	7	US-11-218-305-17633	Sequence 17633, A
10	226.5	19.5	718	6	US-10-953-349-10521	Sequence 10521, A
11	219.5	18.9	640	7	US-11-218-305-17634	Sequence 17634, A
12	203.5	17.6	841	8	US-11-216-545-4173	Sequence 4173, App
13	202.5	17.5	799	6	US-10-953-349-9569	Sequence 9569, App
14	202.5	17.5	828	7	US-11-218-305-17632	Sequence 17632, A
15	202.5	17.5	1023	8	US-11-216-545-4174	Sequence 4174, App
16	183	15.8	1237	8	US-11-216-545-8381	Sequence 8381, App
17	150.5	13.0	882	8	US-11-266-748A-386894	Sequence 386894, A
18	150.5	13.0	882	8	US-11-266-748A-388679	Sequence 388679, A
19	150.5	13.0	882	8	US-11-266-748A-450273	Sequence 450273, A
20	144.5	12.5	947	6	US-10-449-902-24132	Sequence 24132, A
21	141	12.2	627	8	US-11-216-545-6801	Sequence 6801, App
22	136.5	11.8	845	6	US-10-449-902-19818	Sequence 19818, A
23	136.5	11.8	1999	6	US-10-449-902-17846	Sequence 17846, A
24	132.5	11.4	830	6	US-10-449-902-4174	Sequence 4174, App
25	128.5	11.1	2409	6	US-10-449-902-20231	Sequence 20231, A
26	128	11.0	899	6	US-10-449-902-18147	Sequence 18147, A
27	112	9.7	416	6	US-10-523-290-17	Sequence 17, Appl
28	112	9.7	5600	6	US-10-525-318-13	Sequence 13, Appl
29	112	9.7	7000	6	US-10-525-318-8	Sequence 8, Appl
30	88	7.6	1484	7	US-11-218-305-6032	Sequence 6032, App
31	87.5	7.5	1144	8	US-11-266-748A-368279	Sequence 368279, A
32	87.5	7.5	1144	8	US-11-266-748A-451658	Sequence 451658, A
33	87.5	7.5	7320	8	US-11-266-748A-32248	Sequence 32248, A
34	86.5	7.5	2261	6	US-10-449-902-19756	Sequence 19756, A
35	86.5	7.5	3473	7	US-11-218-305-14808	Sequence 14808, A
36	86	7.4	1796	6	US-11-218-305-6036	Sequence 6036, App
37	85.5	7.4	2907	8	US-11-293-697-1067	Sequence 1067, App
38	85.5	7.4	4218	8	US-11-266-748A-28461	Sequence 28461, A
39	85	7.3	2170	7	US-11-218-305-12983	Sequence 12983, A
40	84	7.2	1000	8	US-11-266-748A-295512	Sequence 295512, A
41	84	7.2	1738	6	US-10-196-749-239	Sequence 239, App
42	84	7.2	901	6	US-10-953-349-18012	Sequence 18012, A
43	83.5	7.2	1510	6	US-10-953-349-14286	Sequence 14286, A
44	83.5	7.2	3417	6	US-10-449-902-12992	Sequence 12992, A
45	83.5	7.2				

## ALIGNMENTS

RESULT 1  
 US-11-247-587-124  
 ; Sequence 124, Application US/11247587  
 ; Publication No. US20060130176A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; CURRENT FILING DATE: 2005-10-12  
 ; PRIOR APPLICATION NUMBER: US 60/671,089  
 ; PRIOR FILING DATE: 2005-04-14  
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
 ; PRIOR FILING DATE: 2004-10-12  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
 ; SEQ ID NO 124  
 ; LENGTH: 781  
 ; TYPE: DNA  
 ; ORGANISM: Nicotiana benthamiana  
 ; US-11-247-587-124

Alignment Scores:  
 Pred. No.: 1.39e-22 Length: 781  
 Score: 281.50 Matches: 80  
 Percent Similarity: 58.3% Conservative: 22

Best Local Similarity: 45.7% Mismatches: 46  
 Query Match: 24.3% Indels: 27  
 DB: 8 Gaps: 7

US-10-718-311-8 (1-227) x US-11-247-587-124 (1-781)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 39 ATGGCTTCTCAGTCTTCTCCTCAGCAGAGTGGCCCGGAGAAATGCTCAAGCT 98  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 99 AACATGGTTGCACCTTTCACAGGCTTAAGTCTGTGCTCCTCATCCCTGTTTCAAGAAAG 158  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60  
 DB 159 CAAACCTTGCATCATCTCCATTGCCAGCAACGGCGGAAGAGTGCATGTCAGGTG 218  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 DB 219 TGGCCACCAATTAAC-----ATGAAGAAGTATGAGACTCTCTCATACCTTCCCGATTG 272  
 QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 DB 273 AGC-----CAGGAGCAATTTGCTCCGAAATTTAGTACCTTTTGAAAAAATGGATGGTT 326  
 QY 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118  
 DB 327 CCTTCTGCTGGAAATTCGAG-----ACTGAGAAAGGATTGTC 362  
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 DB 363 TACCGTGAACACCAAGTCACCGAGTACTAT-----GATGGCAGATAC 407  
 QY 137 TrpLeuArgGluIleLeu-LeuCysAlaAspGly-----GluProTrpLeu-A 152  
 DB 408 TGGACCATGTGGAAAGCTACTATGTTGGATGCACTGATGCCACCAAGTGTGGCTGAG 467  
 QY 152 laGlyArgThrValValProValSerThrLeuSerGlyPro 165  
 DB 468 GTGGGAGAGCGGAAGAAGATACCACAGCGCTGGGTCCG 508

RESULT 2

US-11-247-587-125  
 ; Sequence 125, Application US/11247587  
 ; Publication No. US20060130176A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; CURRENT FILING DATE: 2005-10-12  
 ; PRIOR APPLICATION NUMBER: US 60/671,089  
 ; PRIOR FILING DATE: 2005-04-14  
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
 ; PRIOR FILING DATE: 2004-10-12  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
 ; SEQ ID NO 125  
 ; TYPE: DNA  
 ; LENGTH: 762  
 ; ORGANISM: Nicotiana benthamiana  
 US-11-247-587-125

Alignment Scores:  
 Pred. No.: 4,948-22 Length: 762  
 Score: 276.50 Matches: 74

Percent Similarity: 60.8% Conservative: 19  
 Best Local Similarity: 48.4% Mismatches: 34  
 Query Match: 23.9% Indels: 26  
 DB: 8 Gaps: 7

US-10-718-311-8 (1-227) x US-11-247-587-125 (1-762)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 27 ATGGCTTCTCAGTCTTCTCCTCAGCAGAGTGGCCCGGAGAAATGCTCAAGCT 86  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 87 AACATGGTTGCACCTTTCACAGGCTTAAGTCTGTGCTCCTCATCCCTGTTTCAAGAAAG 146  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60  
 DB 147 CAAACCTTGCATCATCTCCATTGCCAGCAACGGCGGAAGAGTGCATGTCAGGTG 206  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArg-----AlaLeuArgTyrCysLys 78  
 DB 207 TGG-----CCACCAATTAACAGAAGAGTACGAGACTCTCTCATACCTTCTCT 254  
 QY 79 GluIleProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSer 96  
 DB 255 GATCTGAGCGTG-----GAGCAATTTGCTTAGCGAAATTTAGTACCTCTTGAAAAAATGGA 308  
 QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
 DB 309 TGGGTTCCTTCTGGAAATTCGAG-----ACTGAGCGCGGA 344  
 QY 117 PheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSer 134  
 DB 345 TTGTGTACCTGACACCAAGTACCGCGGATACTAT-----GACGGC 389  
 QY 135 ArgTyrTrpLeuArgGluIleLeu-LeuCysAlaAsp 146  
 DB 390 AGATACTGGACCATGTGGAAAGTGTGCTATGTTGGAT 426

RESULT 3

US-11-247-587-128  
 ; Sequence 128, Application US/11247587  
 ; Publication No. US20060130176A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Rockefeller University  
 ; APPLICANT: Toboda, Jose Reyes  
 ; APPLICANT: Zhang, Xiuren  
 ; APPLICANT: Soyano, Takashi  
 ; APPLICANT: Chua, Nam-Hai  
 ; APPLICANT: Niu, Qi-Wen  
 ; APPLICANT: Lin, Shih-Shun  
 ; TITLE OF INVENTION: MicroRNAs  
 ; FILE REFERENCE: 2312-124  
 ; CURRENT APPLICATION NUMBER: US/11/247,587  
 ; CURRENT FILING DATE: 2005-10-12  
 ; PRIOR APPLICATION NUMBER: US 60/671,089  
 ; PRIOR FILING DATE: 2005-04-14  
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
 ; PRIOR FILING DATE: 2004-10-12  
 ; NUMBER OF SEQ ID NOS: 174  
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
 ; SEQ ID NO 128  
 ; LENGTH: 698  
 ; TYPE: DNA  
 ; ORGANISM: Nicotiana benthamiana  
 US-11-247-587-128

Alignment Scores:  
 Pred. No.: 1,768-19 Length: 698  
 Score: 253.50 Matches: 72  
 Percent Similarity: 52.7% Conservative: 16  
 Best Local Similarity: 43.1% Mismatches: 40  
 Query Match: 21.9% Indels: 39

DB: 8 Gaps: 8  
US-10-718-311-8 (1-227) x US-11-247-587-128 (1-698)  
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaAlaThrArgSerAsnValThrGln 19  
Db 20 ATGGCTTCCTCTGTGATTCCTCAGCTGTGCGGTGCCACCGCGCTAATGCTGCTCAA 79  
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39  
Db 80 GCAGCATGTTGCACCCCTTCACTGGCCTCAAAATCTGCTTCCTCTCCCTGTTACCAGA 139  
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59  
Db 140 AAACAAACCTTGACATTCATCCATGCTAGCAATGGTGGAGAGTCCAAATGCATGCAG 199  
QY 60 ValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGlu 79  
Db 200 GTGTGCCCAACCAATTAAAC-----ATGAAGAAGTACGACACACTCTCATACCTTCCTGAT 253  
QY 80 IleProAlaLeuAspProGlnLeuLeuAsp-----Trp 90  
Db 254 TTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTATCTTTTGAATAATGATGG 307  
QY 91 LeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrValSerVal 110  
Db 308 GTTCCTTGTGGAATTCGAGACTGAGCGT-----337  
QY 111 ThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuPro 128  
Db 338 -----GGATTTGCTACCGTGAACATCAGCTACACAGGATACTAC---379  
QY 129 LeuLeuProLysGluSerArgTyrTrpLeuArg-----GluLeuLeuCysAla 145  
Db 380 -----GATGGCAGATACCTGACCATGTGGAAGTTGCCCATGTTCGGGTGCACT 427  
QY 146 AspGlyGluProTrpLeuAla 152  
Db 428 GATGCCACTCAGGTGTGGCT 448

## RESULT 4

US-11-247-587-126  
; Sequence 126, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 126  
; LENGTH: 683  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana  
US-11-247-587-126

Alignment Scores:  
Pred. No.: 8,14e-19 Length: 683  
Score: 247.50 Matches: 70  
Percent Similarity: 52.7% Conservative: 18  
Best Local Similarity: 41.9% Mismatches: 40

Query Match: 21.4% Indels: 39  
DB: 8 Gaps: 8  
US-10-718-311-8 (1-227) x US-11-247-587-126 (1-683)  
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaAlaThrArgSerAsnValThrGln 19  
Db 37 ATGGCTTCCTCTGTGATTCCTCAGCTGTGCGGTGCCACCGCGCTAATGCTGCTCAA 96  
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39  
Db 97 GCACATGTTGCACCCCTTCACTGGCCTCAAGTCGCTCCTCTCCCTGTTACCAG 156  
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59  
Db 157 AAACAAACCTTGACATTCATCCATGCTAGCAATGGTGGAGAGTCCAAATGCATGCAG 216  
QY 60 ValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGlu 79  
Db 217 GTGTGCCCAACCAATTAAAC-----ATGAAGAAGTACGACACACTCTCATACCTTCCTGAT 270  
QY 80 IleProAlaLeuAspProGlnLeuLeuAsp-----Trp 90  
Db 271 TTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTATCTTTTGAATAATGATGG 324  
QY 91 LeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrValSerVal 110  
Db 325 GTTCCTTGTGGAATTCGAGACTGAGCGT-----354  
QY 111 ThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuPro 128  
Db 355 -----GGATTTGCTACCGTGAACATCAGCTACACAGGATACTAC---396  
QY 129 LeuLeuProLysGluSerArgTyrTrpLeuArg-----GluLeuLeuCysAla 145  
Db 397 -----GATGGCAGATACCTGACCATGTGGAAGTTGCCCATGTTCGGGTGCACT 444  
QY 146 AspGlyGluProTrpLeuAla 152  
Db 445 GATGCCACTCAGGTGTGGCT 465

## RESULT 5

US-11-247-587-127  
; Sequence 127, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 127  
; LENGTH: 700  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana  
US-11-247-587-127

Alignment Scores:  
Pred. No.: 1.1e-18 Length: 700  
Score: 246.50 Matches: 71  
Percent Similarity: 52.1% Conservative: 17

Best Local Similarity: 42.0% Mismatches: 38  
Query Match: 21.3% Indels: 43  
DB: 8 Gaps: 9

US-10-718-311-8 (1-227) x US-11-247-587-127 (1-700)

QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19  
DB 2 ATGGCTTCTCAGTTATGCTCTCAGCTGCCGCTGTTCACCGGGCCCAATGCTCAA 61  
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39  
DB 62 GCCAGTAGGTGCACTTCACTGGCCTCAAGTCGCAACCTCTCCCTGTTCCAGA 121  
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59  
DB 122 AAACAAACCTTGACATTTACTTCCATTCCTAGCAACGGGGAAGTAGTTCATGATGCG 181  
QY 60 ValTrpHisMetSerHisProAlaLeuThrGlnLeuArg-----AlaLeuArgTyrCys 77  
DB 182 GTGTGG-----CCCAATTAAACAAGAAAGTAGAGACACTCTCATACCTT 229  
QY 78 LysGluIleProAlaLeuAspProGlnLeuLeuAsp----- 89  
DB 230 CCCGATTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTACCTGTGAAAAAT 283  
QY 90 ---TrpLeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrVal 108  
DB 284 GGATGGGTCTCTGCTGGAATTCAGACTGAGCGT----- 319  
QY 109 SerValThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGlu 126  
DB 320 -----GGATTCTGTACCGTGAACACACACAGCTCACCAGGATAT 358  
QY 127 LeuProLeuLeuProLysGluSerArgTyrTrpLeuArg-----GluIleLeuLeu 143  
DB 359 TAT-----GATGGCAGATACCTGGACCATGTGGAAAGTTGCCCATGTTCGGG 403  
QY 144 CysAlaAspGlyGluProTrpLeuAla 152  
DB 404 TGCACGTGATGCCACTCAGGTGTGGCT 430

## RESULT 6

US-11-218-305-17631  
; Sequence 17631, Application US/11218305  
; Publication No. US20060141495A1  
; GENERAL INFORMATION:  
; APPLICANT: MONSANTO TECHNOLOGY, LLC  
; APPLICANT: McIaird, Paul L.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Wu, Kunsheng  
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping  
; FILE REFERENCE: 38-21 (53660)B  
; CURRENT APPLICATION NUMBER: US/11/218,305  
; PRIOR FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US 60/606,880  
; PRIOR FILING DATE: 2004-09-01  
; NUMBER OF SEQ ID NOS: 25043  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17631  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Zea mays

US-11-218-305-17631

## Alignment Scores:

Pred. No.: 6.77e-18 Length: 583  
Score: 238.50 Matches: 85  
Percent Similarity: 49.5% Conservative: 20  
Best Local Similarity: 40.1% Mismatches: 47  
Query Match: 20.6% Indels: 62  
DB: 7 Gaps: 11

US-10-718-311-8 (1-227) x US-11-218-305-17631 (1-583)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
DB 33 ATGGCTTCTCAGTTATGCTCTCCTCCGCGCTGTGGCTACC-----TCACCGGCTCAAGCC 86  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
DB 87 ACCATGGTGTCTCATTACCGGCTTGAAAGTCATCCGCTGCATTCCTCCAGTACCCGCCAAA 146  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
DB 147 AGCAACACTGATATTACTTCCATTGCAACGAGGAGGAGGAGTAGCTGTGATGAAGGTG 206  
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
DB 207 TGGCCACCAAGTCG-----GAAAGAAAGATTGTGAGACT-CTCTCTTATCT 250  
QY 80 eProAlaLeu---AspProGlnLeu-----LeuAspTrpLeuLeuLeuGluAspSe 96  
DB 251 TCCTGACCTTACTGACGTTGAATTGGCCAAAGAGTTGACTACCTTCTCCGCAACAAGTG 310  
QY 96 rMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGln 115  
DB 311 GATTCCCTGTGTGTAATTGTAG-----TTGGAGCA 340  
QY 115 uGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuProLysGln 133  
DB 341 CGGATTTGTGTACCGTGAGCATGGAAGCACACCTGGATAC-----GA 385  
QY 133 userArgTyrTrp----- 137  
DB 386 TGGCGGTTATTGGCAATGTGGAAGCTTCTTTGTCGATGCACTGACTCTGCTCAAGT 445  
QY 138 -LeuArgGluIleLeuLeuCysAlaAspGlyGluPro-----Tr 150  
DB 446 GTTGAAGGAGTGAAGAGTGCAGGAGTACCTTAACGCCCTTTATTAGAATCATTT 504  
QY 150 pLeuAlaGlyArgThrValValProValSerThrLeuSerGlyProGluLeuAlaLeuGln 170  
DB 505 GATTTCGACAAACACCG-----TCAGTCCAGTGCATCAGTTTCA 543  
QY 170 nLysLeuGlyLysThrProLeuGlyArgTyrLeu 181  
DB 544 TCGCCTACAGCCACCAAGCTTTCACCGGTCGCTT 577

## RESULT 7

US-11-247-587-129  
; Sequence 129, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 129  
; LENGTH: 727  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana



```
Qy 79 GlutLeuProAlaLeuAsp---ProGlnLeuLeuAspTrpLeuLeuGluAspSerMet 97
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 267 GACCTTACCGAAGTTGAATTGGTAAGGAAGTCGACTACCTTCTCCGCAACAAGTGGATT 326
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 98 ---ThrLysArgPheGluGlnGlnGlyThrValSerValThrMetIleArgGluGly 116
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 327 CTTGTGTGAATTCGAG-----TTGAGGACCGGA 356
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 117 PheVal-----GluGlnAsnGluIleProGluGluLeuProLeuProLysGluSer 134
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 357 TTTGTGTACCGTGAAGCAGGAGACCCCGGATACTAC-----GATGGC 401
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 135 ArgTyrTrp-----Leu 138
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 402 CTTTACTGGACATGTGAAGTTCCTTGTTCGGATGCATGACTCTGCTCAAGTGTG 461
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 139 ArgGluIleLeuLeuCysAlaAspGlyGluPro 149
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 462 AAGGAAGTCCAAGATGCAAAACGGAGTACCT 494
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
```

## RESULT 10

```
US-10-953-349-10521
; Sequence 10521, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10521
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10521
```

```
Alignment Scores:
Pred. No.: 2,1e-16 Length: 718
Score: 226.50 Matches: 73
Percent Similarity: 51.1% Conservative: 19
Best Local Similarity: 40.6% Mismatches: 42
Query Match: 19.5% Indels: 47
DB: 6 Gaps: 9
```

US-10-718-311-8 (1-227) x US-10-953-349-10521 (1-718)

```
Qy 1 MetAlaSerSerValIleSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 32 ATGGCTTCTCTATGCTCTCTCTCGCGCTGGTTACC-----TCCCGGCTCAAGCC 85
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 86 ACCATGGTCGCTCCATTCACCTGGTTGAAGTCATCCGCTCTCTTCCCGGTCAACCGCAAG 145
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 146 GCCAACAAACGACATTACTTCCATCAACAAATGGGGGAAGAGTTAGCTGCATGAAGTGT 205
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluI1 80
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 206 TGGCACCACCATCG-----GAAAGAGAAGAGTTGAGACT-CTATCTTACTT 249
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 80 eProAlaLeu---AspProGlnLeu-----LeuAspTrpLeuLeuGluAspSe 96
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 250 CCTGACCTTACTGACGTCGAATTGGCTAAGGAAGTTGACTACCTTCTCCGCAACAAGTG 309
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 96 rMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgG1 115
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 310 GATTCCTGTGTGAATTCGAG-----TTGGAGCA 339
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
```

```
Qy 115 uGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysG1 133
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 340 CGGATTTGTGTACCGTGAAGCAGGAAACACATCCCGGATACTAC-----GA 384
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 133 userArgTyrTrp----- 137
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 385 TGGACGGTACTGGACAATGTGGAAGCTTCCATTGTTCGGATGCACCGACTCTGCTCAAGT 444
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 138 -LeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 445 ATTGAAGGAAGTTGAAGAATGCAAGAGAGTACCCGGGCGCTTCAATAGGATCATC 502
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
```

## RESULT 11

```
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634
```

```
Alignment Scores:
Pred. No.: 1,1e-15 Length: 640
Score: 219.50 Matches: 65
Percent Similarity: 49.7% Conservative: 20
Best Local Similarity: 38.0% Mismatches: 43
Query Match: 18.9% Indels: 43
DB: 7 Gaps: 9
```

US-10-718-311-8 (1-227) x US-11-218-305-17634 (1-640)

```
Qy 1 MetAlaSerSerValIleSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 33 ATGGCTTACTCTATGCTCTCTCTCGCGCTGTTGTACC-----TCCCGGCTCAAGCC 86
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 87 ACCATGGTCGCTCCATTCACCGGCTTGAAGTCTTCTCTGCAATCCAGTCACCGCAAG 146
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 147 GCCAACAAACGACATTACTTCCATCGTATAGCAACGGAAGAGTTAGCTGCATGAGGTTG 206
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 207 TGG-----CCACCAATTGGAAGAGAGTTTGAGACCTCTCTTACCTCTCT 254
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 79 GlutLeuProAlaLeuAsp---ProGlnLeuLeuAspTrpLeuLeuGluAspSerMet 97
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 255 GACCTTACCGAAGTTGAATTGGTAAGGAAGTTGACTACCTTCTCCGCAACAAGTGGATT 314
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Qy 98 ---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
```



Db 315 CCTGTGTTCAATTCAG-----TTGGAGCAGGA 344  
QY 117 pheVal-----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSer 134  
Db 345 TTTGTGTACCGTGAGCAGCAGGACCCCGGATAC-----GATGGC 389  
QY 135 ArgTyrTrp-----Leu 138  
Db 390 CGTTACTGACAAATGGGAAGCTTCCTTGTTCGGATGCACTCTCTCTCAAGTGTG 449  
QY 139 ArgGluLeuLeuCysAlaaspGlyGluPro 149  
Db 450 AAGGAAGTCCAAGATGCAAACTGATGACCT 482

## RESULT 12

US-11-216-545-4173  
; Sequence 4173, Application US/11216545  
; Publication No. US20060135758A1  
; GENERAL INFORMATION:  
; APPLICANT: MONSANTO Technology, LLC  
; APPLICANT: Mc Laird, Paul L  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Wu, Kunsheng  
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping  
; FILE REFERENCE: 38-21 (53659)B  
; CURRENT APPLICATION NUMBER: US/11/216,545  
; PRIOR FILING DATE: 2005-08-31  
; PRIOR APPLICATION NUMBER: US 60/606,062  
; PRIOR FILING DATE: 2004-08-31  
; NUMBER OF SEQ ID NOS: 8783  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4173  
; LENGTH: 841  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-11-216-545-4173

Alignment Scores:  
Pred. No.: 1,07e-13 Length: 841  
Score: 203.50 Matches: 64  
Percent Similarity: 53.1% Conservative: 13  
Best Local Similarity: 44.1% Mismatches: 37  
Query Match: 17.6% Indels: 31  
DB: 8 Gaps: 10

US-10-718-311-8 (1-227) x US-11-216-545-4173 (1-841)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 37 ATGGCTTCTCAATGATCTCTCCCGAGCTGTACCACC---GTCAACCGTCCGGTGCC 93  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 94 GCGATGTTGCTCCATTCACCGGCTCAATCCATCGCTCGCTTCC---ACGAGGAAG 150  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
Db 151 ACCAACATACATTAACCTCCATTCGTAGCAACGGTGAAGAGTACAATGCATGCAGGTG 210  
QY 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78  
Db 211 TGG-----CCACAATTTGGCAAGAGTTTCGAGACTCTTCTCTACTTGCCA 258  
QY 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95  
Db 259 GACCTCGAT-----GATGCCAATTGGCAAGAGTGAATACCTCTTAAGGAAGGA 312  
QY 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114  
Db 313 TGGATTCTTGTGTTGAATTCGAG-----TTGGAG 342  
QY 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132

Db 343 CACGGTTTCGTACCGTGAGCACACAGGTCACTGGATACTAT----- 387  
QY 133 GluSerArgTyrTrp 137  
Db 388 GATGACGCTACTGG 402  
RESULT 13  
US-10-953-349-9569  
; Sequence 9569, Application US/10953349  
; Publication No. US20060107345A1  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1579PUS2  
; CURRENT APPLICATION NUMBER: US/10/953,349  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 40252  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 9569  
; LENGTH: 799  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-953-349-9569  
Alignment Scores:  
Pred. No.: 1.29e-13 Length: 799  
Score: 202.50 Matches: 58  
Percent Similarity: 53.1% Conservative: 19  
Best Local Similarity: 40.0% Mismatches: 37  
Query Match: 17.5% Indels: 31  
DB: 6 Gaps: 9

US-10-718-311-8 (1-227) x US-10-953-349-9569 (1-799)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 47 ATGGCTTCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCGGCTCAGGCC 100  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 101 ACTATGGTCGCTCTCTTCAACGGACTTAAGTCTCGCTCGCTTCCAGCACCCGCAAG 160  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
Db 161 GCTAACACACGACATTACTTCCATCACAGCAACGGCGGAAGAGTTAACTGCATGCAGGTG 220  
QY 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78  
Db 221 TGG-----CCTCCGATTGGAAAGAGTTTGGAGACTCTCTCTTACCTTCTCT 268  
QY 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95  
Db 269 GACCTT-----ACCGATTCCGAATTTGGCTAAGGAAGTTGACTACCTTATCCGCAACAG 322  
QY 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114  
Db 323 TGGATTCTTGTGTTGAATTCGAG-----TTGGAG 352  
QY 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132  
Db 353 CACGATTGTGTACCGTGACCGTAACCTCACCCTGATCTAT----- 397  
QY 133 GluSerArgTyrTrp 137  
Db 398 GATGACGCTACTGG 412  
RESULT 14  
US-11-218-305-17632  
; Sequence 17632, Application US/11218305  
; Publication No. US20060141495A1  
; GENERAL INFORMATION:

```
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLeaird, Paul L.
; APPLICANT: Tao, Nengsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17632
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17632

Alignment Scores:
Pred. No.: 1.36e-13 Length: 828
Score: 202.50 Matches: 58
Percent Similarity: 53.1% Conservative: 19
Best Local Similarity: 40.0% Mismatches: 37
Query Match: 17.5% Indels: 31
DB: 7 Gaps: 9

US-10-718-311-8 (1-227) x US-11-218-305-17632 (1-828)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 ATGGCTTCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 103
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 ACTATGGTCGCTCTTCAACGGACTTAAGTCTCCGCTGCTTCCCGACCCACCGCAAG 163
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 GCTAACACGACATTACTTCCATCAACGACGCGGAGAGTTAAGTCTGATGTCAGGTG 223
Qy 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 TGG-----CCTCCGATTGGAAGAAGATTGTGAGACTCTCTCTTACCTTCT 271
Qy 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuLeuGluAsp 95
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 GACCTT-----ACCGATTCCGAATTGGCTAAGGAAGTTGACTACCTTTATCCGCAACAA 325
Qy 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 TGGATTCTTGTGTGAATTTCGAG-----TTGGAG 355
Qy 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 CACGGATTGTGTACCGTGACGACCGGTAACTCAACCGGATACAT----- 400
Qy 133 GluSerArgTyrTrp 137
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 GATGGACGGTACTGG 415

RESULT 15
US-11-216-545-4174
; Sequence 4174, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLeaird, Paul L
; APPLICANT: Tao, Nengsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
```

```
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4174
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)-(12)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-4174

Alignment Scores:
Pred. No.: 1.85e-13 Length: 1023
Score: 202.50 Matches: 64
Percent Similarity: 53.1% Conservative: 13
Best Local Similarity: 44.1% Mismatches: 37
Query Match: 17.5% Indels: 31
DB: 8 Gaps: 10

US-10-718-311-8 (1-227) x US-11-216-545-4174 (1-1023)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 ATGGCTTCTCTCAATGACTCTTCCCGAGCTGTACCACCT---GTCAACCGTCCGGTGCC 284
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 GGCAATGGTTGCTTCATTCACTGGCTCAAGTCCATGGCTGGCTTCCCC---ACCAGGAAG 341
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
342 ACCAACAAATGACATTACCTCCATTGCTAGCAACGTTGAAGAGTGCAATGTCATGCAGGTG 401
Qy 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 TGG-----CCACCGATTGGCAAGAAGATTGTGAGACTCTTCTTACCTGCCA 449
Qy 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuLeuGluAsp 95
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 GACCTTGAT-----GATGCACAATTGCAAAAGGAAGTAGAATACTTCTTAGGAAGGA 503
Qy 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 TGGATTCTTGTGTAATTTCGAG-----TTGGAG 533
Qy 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
534 CACGGTTTGTGTACCGTGACGACACACAGGTCAACAGGATCACCAGGATACTAC----- 578
Qy 133 GluSerArgTyrTrp 137
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
579 GATGGACGGTACTGG 593
```

Search completed: July 11, 2006, 21:05:58  
Job time : 137.1 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 20:23:26 ; Search time 1856.55 Seconds  
(without alignments)  
2525.695 Million cell updates/sec

Title: US-10-718-311-8  
Perfect score: 1159  
Sequence: 1 MASSVISAANVATNSVTAQ.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-DB=abs/ABSSWEB spool/US10718311/runat 11072006 11531 5371/app query.fasta\_1  
-WTR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NOR=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808  
-USER=US10718311@CN 1 1.2326 @runat 11072006 11531 5371 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	1159	100.0	684	3 US-09-855-341-7 Sequence 7, Appli

2	1159	100.0	684	3 US-09-896-8668-16 Sequence 16, Appl
3	1159	100.0	684	7 US-10-359-369-41 Sequence 41, Appl
4	1159	100.0	684	8 US-10-699-050-16 Sequence 16, Appl
5	1159	100.0	684	8 US-10-718-311-7 Sequence 7, Appl
6	1159	100.0	684	9 US-10-462-162-33 Sequence 33, Appl
7	862	74.4	3452	9 US-10-462-162-47 Sequence 47, Appl
8	854	73.7	495	3 US-09-855-341-3 Sequence 3, Appl
9	854	73.7	495	3 US-09-896-8668-12 Sequence 12, Appl
10	854	73.7	495	3 US-10-359-369-37 Sequence 37, Appl
11	854	73.7	495	8 US-10-699-050-12 Sequence 12, Appl
12	854	73.7	495	8 US-10-718-311-3 Sequence 3, Appl
13	854	73.7	495	9 US-10-462-162-29 Sequence 29, Appl
14	851	73.4	1971	9 US-10-462-162-54 Sequence 54, Appl
15	813	70.1	1207	10 US-10-450-763-26139 Sequence 26139, A
16	277.5	23.9	489	10 US-10-487-901-12707 Sequence 7207, Ap
17	276.5	23.9	504	10 US-10-487-901-3503 Sequence 3503, Ap
18	276.5	23.9	718	10 US-10-487-901-5765 Sequence 5765, Ap
19	276.5	23.9	736	10 US-10-487-901-1704 Sequence 1704, Ap
20	276.5	23.9	736	10 US-10-487-901-5043 Sequence 5043, Ap
21	269	23.2	204	3 US-09-839-477-3 Sequence 3, Appl
22	269	23.2	204	8 US-10-758-064-3 Sequence 3, Appl
23	269	23.2	297	10 US-10-487-901-7209 Sequence 7209, Ap
24	269	23.2	704	9 US-10-425-115-85509 Sequence 85509, A
25	267.5	23.1	704	9 US-10-425-115-85509 Sequence 85509, A
26	266	23.0	668	10 US-10-487-901-1713 Sequence 1713, Ap
27	253.5	21.9	683	10 US-10-487-901-5045 Sequence 5045, Ap
28	253.5	21.9	737	10 US-10-487-901-5766 Sequence 5766, Ap
29	253.5	21.9	737	10 US-10-487-901-5766 Sequence 5766, Ap
30	247.5	21.4	738	10 US-10-487-901-1716 Sequence 1716, Ap
31	247.5	21.4	753	10 US-10-487-901-5764 Sequence 5764, Ap
32	247.5	21.4	754	10 US-10-487-901-5028 Sequence 5028, Ap
33	246.5	21.3	714	10 US-10-487-901-1703 Sequence 1703, Ap
34	246.5	21.3	736	10 US-10-487-901-5044 Sequence 5044, Ap
35	244	21.1	174	3 US-09-854-286-15 Sequence 15, Appl
36	236.5	20.4	654	10 US-10-487-901-5042 Sequence 5042, Ap
37	235	20.3	169	6 US-10-165-420-3 Sequence 3, Appl
38	230.5	19.9	629	10 US-10-487-901-225 Sequence 225, App
39	228.5	19.7	724	3 US-09-910-664-55 Sequence 55, Appl
40	228.5	19.7	724	8 US-10-333-184-54 Sequence 54, Appl
41	228.5	19.7	726	8 US-10-333-184-202 Sequence 202, App
42	228.5	19.7	764	10 US-10-487-901-4119 Sequence 4119, Ap
43	228.5	19.7	771	3 US-09-770-445-918 Sequence 918, App
44	225.5	19.5	1205	3 US-09-770-445-41 Sequence 41, Appl
45	225	19.4	385	3 US-09-732-627A-1176 Sequence 1176, Ap

ALIGNMENTS

RESULT 1  
US-09-855-341-7  
; Sequence 7, Application US/09855341  
; Patent No. US2002002715A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; TITLE OF INVENTION: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL  
US-09-855-341-7

Alignment Scores:	6.06e-139	Length:	684
Pred. No.:			



; APPLICANT: Viitanen, Paul  
; APPLICANT: Meyer, Knut  
; APPLICANT: Van Dyk, Drew  
; TITLE OF INVENTION: UDP-Glucosyltransferases  
; FILE REFERENCE: CL1821 US NA  
; CURRENT APPLICATION NUMBER: US/10/359,369  
; CURRENT FILING DATE: 2003-02-06  
; PRIOR APPLICATION NUMBER: 60/355,511  
; PRIOR FILING DATE: 2002-02-07  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 41  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(684)  
US-10-359-369-41

Alignment Scores:  
Pred. No.: 6,06e-139 Length: 684  
Score: 1159.00 Matches: 227  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 7 Gaps: 0

US-10-718-311-8 (1-227) x US-10-359-369-41 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
DB 1 ATGGCTTCCTCTGTCTATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAGCT 60  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
DB 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCCACTTCCTCTGTACAAAGAG 120  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
DB 121 CAAAACCTTGACATCCTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAAGTG 180  
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
DB 181 TGGCATATGTACACCCCGCTTACGCAACTGCGTGGCGCTGCGCTATTGTAAAGAGATC 240  
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100  
DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTTCGGAGGATTCATGACAAAACGT 300  
QY 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
DB 301 TTGTGAACAGCAGGAGAAACCGGTAAAGCGTGACATGATCCGCGAAGGGTTTGTGAGCAG 360  
QY 121 AsnGluIleProGluLeuProLeuLysGluSerArgTyrTrpLeuArgGlu 140  
DB 361 AATGAATCCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420  
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160  
DB 421 ATTTTGTATGTGCCATGGTGAACCGTGGCTTGGCGGTCTGACCGTCTCTGTGTCA 480  
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180  
DB 481 ACGTTAAGCGGCGCGAGCTGGCGTTTACAAAATTGGGTAAACCGCGTTAGACGCTAT 540  
QY 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200  
DB 541 CTGTTACATCATCGACATTAAACCCGGGACTTTATTGAGATAGCGCGTATGCCGGCTG 600  
QY 201 TrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGlnLeuPhe 220  
DB 601 TGGGGCGACGTTCCCGCTCGCATTAAGCGGTAAACCGCTGTTCGTAAACAGACTGTTT 660

QY 221 LeuProAlaSerProLeuTyr 227  
DB 661 TTACCGGCGTCCACCGTTGTAC 681

## RESULT 4

US-10-699-050-16  
; Sequence 16, Application US/10699050  
; Publication No. US20040142437A1  
; GENERAL INFORMATION:  
; APPLICANT: Flint, Dennis  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate  
; FILE REFERENCE: BC1034 US NA  
; CURRENT APPLICATION NUMBER: US/10/699,050  
; CURRENT FILING DATE: 2003-10-30  
; PRIOR APPLICATION NUMBER: US/09/896,866B  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/216,615  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei  
US-10-699-050-16

## Alignment Scores:

Pred. No.: 6,06e-139 Length: 684  
Score: 1159.00 Matches: 227  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 8 Gaps: 0

US-10-718-311-8 (1-227) x US-10-699-050-16 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
DB 1 ATGGCTTCCTCTGTCTATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAGCT 60  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
DB 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCCACTTCCTCTGTACAAAGAG 120  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
DB 121 CAAAACCTTGACATCCTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAAGTG 180  
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
DB 181 TGGCATATGTACACCCCGCTTACGCAACTGCGTGGCGCTGCGCTATTGTAAAGAGATC 240  
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100  
DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTTCGGAGGATTCATGACAAAACGT 300  
QY 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
DB 301 TTGTGAACAGCAGGAGAAACCGGTAAAGCGTGACATGATCCGCGAAGGGTTTGTGAGCAG 360  
QY 121 AsnGluIleProGluLeuProLeuLysGluSerArgTyrTrpLeuArgGlu 140  
DB 361 AATGAATCCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420  
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160

Db 421 ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGGCGGTCTGTACCGTCTGTTCTCTGTGCA 480  
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnIlyLeuGlyLeuThrProLeuGlyArgTyr 180  
Db 481 ACGTTAAGCGGGCGGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540  
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyValArgAspAlaGlyLeu 200  
Db 541 CTGTTTCAATCATCATGACATTAAACCGGACCTTTATTGAGATAGGCGGTGATGCCGGGCTG 600  
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyIlyProLeuLeuLeuThrGluLeuPhe 220  
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTCTTAACAGAACTGTTT 660  
Qy 221 LeuProAlaSerProLeuTyr 227  
Db 661 TTACCGGGCTCACCGTTGTAC 681

## RESULT 5

US-10-718-311-7  
; Sequence 7, Application US/10718311  
; Publication No. US20040143867A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUZ  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; TITLE OF INVENTION: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/10/718,311  
; CURRENT FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: US/09/855,341  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL  
US-10-718-311-7

## Alignment Scores:

Pred. No.: 6,06e-139 Length: 684  
Score: 1159.00 Matches: 227  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 8 Gaps: 0

US-10-718-311-8 (1-227) x US-10-718-311-7 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValIleThrArgSerAsnValThrGlnAla 20  
Db 1 ATGGCTTCTCTGTCTATTTCTTCAGCAGCTGTTGCGACAGCAGCAATGTTACACAAGCT 60  
Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlylys 40  
Db 61 AGCATGGTTGCACCTTTCTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120  
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyIlyArgValSerCysMetGlnVal 60  
Db 121 CAAAACCTTGACATCATTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180  
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysIlysgluIle 80  
Db 181 TGGCATATGTACACCCCGGTTAACGCAACTGCGCTGGCGCTATTGTAAAGAGATC 240  
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrIlylsArg 100  
Db 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATACAAAACGT 300

Qy 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
Db 301 TTTGACACGACGAGGAAAAACGGTAAAGCTGACGATGATCCCGAAGGGTGTTCGAGCAG 360  
Qy 121 AsnGluIleProGluGluLeuProLeuLeuProIlySerGlySerArgTyrTrpLeuArgGlu 140  
Db 361 AATGAAATCCCGAAGAACTGCGCTGTCGCGAAGAGTCTCGTTACTGTTACGTGAA 420  
Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160  
Db 421 ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGGCGGTCTGACCGTCTCTGTGCA 480  
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnIlyLeuGlyLysThrProLeuGlyArgTyr 180  
Db 481 ACGTTAAGCGGGCGGAGCTGCGCTTACAAAATTTGGGTAAACCGCGCTTAGGACGCTAT 540  
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200  
Db 541 CTGTTTCAATCATCATGACATTAAACCGGACCTTTATTGAGATAGGCGGTGATGCCGGGCTG 600  
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyIlyProLeuLeuLeuThrGluLeuPhe 220  
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTCTTAACAGAACTGTTT 660  
Qy 221 LeuProAlaSerProLeuTyr 227  
Db 661 TTACCGGGCTCACCGTTGTAC 681

## RESULT 6

US-10-462-162-33  
; Sequence 33, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA  
; CURRENT APPLICATION NUMBER: US/10/462,162  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 33  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-462-162-33

Alignment Scores:  
Pred. No.: 6,06e-139 Length: 684  
Score: 1159.00 Matches: 227  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-10-718-311-8 (1-227) x US-10-462-162-33 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValIleThrArgSerAsnValThrGlnAla 20  
Db 1 ATGGCTTCTCTGTCTATTTCTTCAGCAGCTGTTGCGACAGCAGCAATGTTACACAAGCT 60  
Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlylys 40  
Db 61 AGCATGGTTGCACCTTTCTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120  
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyIlyArgValSerCysMetGlnVal 60  
Db 121 CAAAACCTTGACATCATTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180  
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysIlysgluIle 80

Db 181 TGGCATATGTCACACCCCGGTTAACGCAACTCGCTGCGCTATTGTTAAAGAGATC 240  
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100  
Db 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGGCTGGAGGATTCCATGACAAACGT 300  
Qy 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
Db 301 TTGAAACAGCAGGAGGAAAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTGAGCAG 360  
Qy 121 AsnGluLeuProGluLeuProLeuProLeuProLysGluSerArgTyrTrpLeuArgGlu 140  
Db 361 ATGAATATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGGTTACGTGAA 420  
Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160  
Db 421 ATTTTGTATGTCGCGATGGTAACCGTGGCTTGCCTGCGTACCGCTGCTGCTGTGTC 480  
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180  
Db 481 ACGTTAAGCGCGCGGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGACGCTAT 540  
Qy 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200  
Db 541 CTGTTACATCATCGACATTAACCGGAGCTTTATTGAGTAGGCGGTGATGCGGGCTG 600  
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220  
Db 601 TGGGGCGGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 660  
Qy 221 LeuProAlaSerProLeuTyr 227  
Db 661 TTACCGGCGCTCACCGGTTGTAC 681

## RESULT 7

US-10-462-162-47  
; Sequence 47, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA  
; CURRENT APPLICATION NUMBER: US/10/462,162  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 47  
; LENGTH: 3452  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of nucleic acid fragment encoding CPL (from  
; OTHER INFORMATION: Escherichia coli), PHBA 1-hydroxylase (from Candida  
; OTHER INFORMATION: parapsilosis), and UGT72B1 (from Arabidopsis thaliana).  
US-10-462-162-47

Alignment Scores:  
Pred. No.: 1 57e-99 Length: 3452  
Score: 862.00 Matches: 166  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 74.4% Indels: 0  
DB: 9 Gaps: 0

US-10-718-311-8 (1-227) x US-10-462-162-47 (1-3452)

Qy 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuPro 81  
Db 1 CATATGTACACCCCGGTTAACGCAACTCGGTGCGCTGCGCTATTGTTAAAGAGATCCCT 60

Qy 82 AlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPhe 101  
Db 61 GCCTCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAACGTTTT 120  
Qy 102 GluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGlnAsn 121  
Db 121 GAACAGCAGGAGGAAAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTGAGCAGAA 180  
Qy 122 GluLeuProGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 141  
Db 181 GAAATCCCGGAAACCTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAAAT 240  
Qy 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThr 161  
Db 241 TTGTTATGTCGCGATGGTAACCGTGGCTTGCCTGCGTACCGTCTCTCTGTGTCAACG 300  
Qy 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181  
Db 301 TTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCTATCTG 360  
Qy 182 PheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTyr 201  
Db 361 TTTACATCATCGACATTAACCGGAGCTTTATTGAGTAGGCGGTGATGCGGGCTGTGG 420  
Qy 202 GlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221  
Db 421 GGGCGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 480  
Qy 222 ProAlaSerProLeuTyr 227  
Db 481 CCGGCGCTCACCGGTTGTAC 498

## RESULT 8

US-09-855-341-3  
; Sequence 3, Application US/09855341  
; Patent No. US20020002715A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; TITLE OF INVENTION: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 3  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:E. coli  
US-09-855-341-3

Alignment Scores:  
Pred. No.: 9 77e-100 Length: 495  
Score: 854.00 Matches: 165  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 73.7% Indels: 0  
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-855-341-3 (1-495)

Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82  
Db 1 ATGTACACCCCGGTTAACGCAACTCGGTGCGCTGCGCTATTGTTAAAGAGATCCCTGCC 60  
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAACCGTTTGA 120



Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGGGAAAACGGTAAGCGTAGCATGATCCCGAAGGGTTGTTCGACGAGATGAA 180  
Qy 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
Db 181 ATCCCGAAGAACTCCGCTGTCGCGAAGAGTCTCGTTACTGCTTACGTGAATTTTG 240  
Qy 143 LeuGlyAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162  
Db 241 TTATGTGCGGATGGTAACCGTGGCTGTCGCGTCTGATCGTTCCTGTGTCAACGTTA 300  
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
Db 301 AGCGGGCGGAGCTGGCTTACAAAAATTGGTAAACGCCGTTAGGAGCTATCTGTTTC 360  
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTAAACCGGACTTTATTGAGATAGGCGCGTGTAGCGGGCTGTGGGG 420  
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222  
Db 421 CGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTTACCG 480  
Qy 223 AlaSerProLeuTyr 227  
Db 481 CGGTACCGTTGTATC 495

## RESULT 9

US-09-896-866B-12

; Sequence 12, Application US/0989866B

; Patent No. US20020151002A1

; GENERAL INFORMATION:

; APPLICANT: Flint, Dennis

; APPLICANT: Meyer, Knut

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Synapoylglucose:Malate Synapoyltransferase Form Malate Conjugates

; FILE REFERENCE: Benzoic Acid Glucosides

; CURRENT APPLICATION NUMBER: US/09/896,866B

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 495

; TYPE: DNA

; ORGANISM: Escherichia coli

US-09-896-866B-12

Alignment Scores:  
Pred. No.: 9,77e-100 Length: 495  
Score: 854.00 Matches: 165  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 73.7% Indels: 0  
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-896-866B-12 (1-495)

Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
Db 1 ATGTACACACCGCGTTAAACGCACTGCGTGGCTGCTATTGTAAAGAGATCCCTGCC 60  
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTGCTCGATGCGTGTGTGGAGGATTCATGACAAAAACGTTTGAA 120  
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGGGAAAACGGTAAGCGTGNACGATGATCCCGAAGGGTTGTTCGACGAGATGAA 180

Qy 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
Db 181 ATCCCGGAGAACTGCGCTGCTCCGNAAGAGTCTCGTTACTGCTTACGTGAATTTTG 240  
Qy 143 LeuGlyAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162  
Db 241 TTATGTGCGGATGGTAACCGTGGCTTACAAAAATTGGTAAACGCCGTTAGGAGCTATCTGTTTC 300  
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
Db 301 AGCGGGCGGAGCTGGCTTACAAAAATTGGTAAACGCCGTTAGGAGCTATCTGTTTC 360  
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTAAACCGGACTTTATTGAGATAGGCGCGTGTAGCGGGCTGTGGGG 420  
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222  
Db 421 CGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTTACCG 480  
Qy 223 AlaSerProLeuTyr 227  
Db 481 CGGTACCGTTGTATC 495

## RESULT 10

US-10-359-369-37

; Sequence 37, Application US/10359369

; Publication No. US20030215927A1

; GENERAL INFORMATION:

; APPLICANT: E. I. duPont de Nemours and Company, Inc.

; APPLICANT: Viitanen, Paul

; APPLICANT: Meyer, Knut

; APPLICANT: Van Dyk, Drew

; TITLE OF INVENTION: UDP-Glucosyltransferases

; FILE REFERENCE: CU1821 US NA

; CURRENT APPLICATION NUMBER: US/10/359,369

; CURRENT FILING DATE: 2003-02-06

; PRIOR APPLICATION NUMBER: 60/355,511

; PRIOR FILING DATE: 2002-02-07

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 37

; LENGTH: 495

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(495)

US-10-359-369-37

Alignment Scores:  
Pred. No.: 9,77e-100 Length: 495  
Score: 854.00 Matches: 165  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 73.7% Indels: 0  
DB: 7 Gaps: 0

US-10-718-311-8 (1-227) x US-10-359-369-37 (1-495)

Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
Db 1 ATGTACACACCGCGTTAAACGCACTGCGTGGCTGCTATTGTAAAGAGATCCCTGCC 60  
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTGCTCGATGCGTGTGTGGAGGATTCATGACAAAAACGTTTGAA 120  
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGGGAAAACGGTAAGCGTGNACGATGATCCCGAAGGGTTGTTCGACGAGATGAA 180  
Qy 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142



Db 181 ATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTG 240  
QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162  
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTCTGTGTCAACGTTA 300  
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgThrLeuPhe 182  
Db 301 AGCGGCCCGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGACGCTATCTGTTTC 360  
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTACCCGGGACCTTTATTGATAGGCGGTGATCCGGCTGTGGGG 420  
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
Db 421 CGACGTTCCTCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480  
QY 223 AlaSerProLeuTyr 227  
Db 481 GCGTCACCGTTGTAC 495

## RESULT 11

US-10-699-050-12

; Sequence 12, Application US/10599050

; Publication No. US20040142437A1

; GENERAL INFORMATION:

; APPLICANT: Meyer, Knut

; APPLICANT: Flint, Dennis

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugates

; FILE REFERENCE: BC1034 US NA

; CURRENT APPLICATION NUMBER: US/10/699,050

; PRIOR FILING DATE: 2003-10-30

; PRIOR APPLICATION NUMBER: US/09/896,866B

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 495

; TYPE: DNA

; ORGANISM: Escherichia coli

US-10-699-050-12

## Alignment Scores:

Pred. No.:	9,77e-100	Length:	495
Score:	854.00	Matches:	165
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	73.7%	Indels:	0
DB:	8	Gaps:	0

US-10-718-311-8 (1-227) x US-10-699-050-12 (1-495)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
Db 1 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTATTGTTAAAGAGATCCCTGCC 60  
QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATGACAAAACGTTTGA 120  
QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGGAAAAACCGTAAACCGTACGATGATCCGGAAGGGTTGTCGAGCAGAAATGA 180  
QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
Db 181 ATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTG 240

QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162  
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTCTGTGTCAACGTTA 300  
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgThrLeuPhe 182  
Db 301 AGCGGCCCGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGACGCTATCTGTTTC 360  
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTACCCGGGACCTTTATTGATAGGCGGTGATCCGGCTGTGGGG 420  
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
Db 421 CGACGTTCCTCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480  
QY 223 AlaSerProLeuTyr 227  
Db 481 GCGTCACCGTTGTAC 495

## RESULT 12

US-10-718-311-3

; Sequence 3, Application US/10718311

; Publication No. US20040143867A1

; GENERAL INFORMATION:

; APPLICANT: VIITANEN, PAUL V.

; APPLICANT: MEYER, KNUT

; APPLICANT: VAN DYK, DREW

; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID

; TITLE OF INVENTION: IN GREEN PLANTS

; FILE REFERENCE: BC1015 US NA

; CURRENT APPLICATION NUMBER: US/10/718,311

; PRIOR FILING DATE: 2003-11-20

; PRIOR APPLICATION NUMBER: US/09/855,341

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: MICROSOFT OFFICE 97

; SEQ ID NO 3

; LENGTH: 495

; TYPE: DNA

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: E. coli

US-10-718-311-3

## Alignment Scores:

Pred. No.:	9,77e-100	Length:	495
Score:	854.00	Matches:	165
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	73.7%	Indels:	0
DB:	8	Gaps:	0

US-10-718-311-8 (1-227) x US-10-718-311-3 (1-495)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
Db 1 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTATTGTTAAAGAGATCCCTGCC 60  
QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATGACAAAACGTTTGA 120  
QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGGAAAAACCGTAAACCGTACGATGATCCGGAAGGGTTGTCGAGCAGAAATGA 180  
QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
Db 181 ATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTG 240  
QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162

```
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTGCCGGTCGTACCGTCTTCTCTGTGTCAACGTTA 300
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 360
Qy 183 ThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCCGTGATGCCGGCTGTGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCCTCGGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTAC 495
RESULT 13
US-10-462-162-29
; Sequence 29, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462.162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-29
Alignment Scores:
Pred. No.: 9,86e-100 Length: 498
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 9 Gaps: 0
US-10-718-311-8 (1-227) x US-10-462-162-29 (1-498)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
Db 1 ATGTGCACACCCCGCTTAACGCNACTGGCGGCTGCCTATTGTAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATCACAAAACGTTTTGAA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGlnAsnGlu 122
Db 121 CAGCAGGGAAAAACGGTAAGCGTAGATATCCCGGAAGGGTTTGTGAGCAGAAATGAA 180
Qy 123 IleProGluLeuLeuProLeuProLysGlySerArgTyrTrpLeuArgGluIleLeu 142
Db 181 ATCCCGGAGAACTGCCGCTGTCGGAAGAGTCTCGTACTGTTAGTGAAATTTTG 240
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTGCCGGTCGTACCGTCTCTGTGTCAACGTTA 300
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 360
```

```
Qy 183 ThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCCGTGATGCCGGCTGTGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCCTCGGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTAC 495
RESULT 14
US-10-462-162-54
; Sequence 54, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462.162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 54
; LENGTH: 1971
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of a nucleic acid fragment inserted into
; OTHER INFORMATION: expression vector pET19a encoding CPL (from Escherichia coli) and
; OTHER INFORMATION: pHBA 1-hydroxylase (from Cadida parapsilosis).
US-10-462-162-54
Alignment Scores:
Pred. No.: 1,82e-98 Length: 1971
Score: 851.00 Matches: 165
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 73.4% Indels: 0
DB: 9 Gaps: 0
US-10-718-311-8 (1-227) x US-10-462-162-54 (1-1971)
Qy 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIlePro 81
Db 1 CATATGTACACCCCGCTTAACGCNACTGGCGTGCCTATTTTAAAGAGATCCCT 60
Qy 82 AlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPhe 101
Db 61 GCCCTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATCACAAAACGTTTT 120
Qy 102 GluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGlnAsn 121
Db 121 GAACAGCAGGAAAAACGGTAAGCGTAGCATGATCCCGGAAGGGTTTGTGAGCAGAAAT 180
Qy 122 GluIleProGluLeuLeuProLeuLeuProLysGlySerArgTyrTrpLeuArgGluIle 141
Db 181 GAAATCCCGAAGAACTGCCGCTGTCGCGAAAGAGTCTCGTTACTGGTTAGTGAATTT 240
Qy 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThr 161
Db 241 TTGTTATGTGCGGATGGTGAACCGTGGCTTGCCGGTCGTACCGTCTCTCTGTGTCAACG 300
Qy 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181
Db 301 TTAAGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTG 360
Qy 182 PheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrp 201
```

Db 361 TTCACATATCGACATTACCCGGGACTTATTGAGATAGGCGGTGATGCCGGGCTGTGG 420  
QY 202 GlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221  
Db 421 GGGCGACGTTCCCGCTCGATTAAACGGTAAACCGCTGTGCTAACACAGACTGTTTTTA 480  
QY 222 ProAlaSerProLeuTyr 227  
Db 481 CCGGCGTCACCGTGTGAC 498

## RESULT 15

US-10-450-763-26139  
; Sequence 26139, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; PRIOR FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 26139  
; LENGTH: 1207  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (893)...(925)  
; OTHER INFORMATION: 81% homologous to Escherichia coli 4-hydroxybenzoate  
; OTHER INFORMATION: synthetase, accession number X57434, Smith-Waterman Score=49.  
US-10-450-763-26139

## Alignment Scores:

Pred. No.:	7,21e-94	Length:	1207
Score:	813.00	Matches:	164
Percent Similarity:	96.0%	Conservative:	2
Best Local Similarity:	94.8%	Mismatches:	5
Query Match:	70.1%	Indels:	2
DB:	10	Gaps:	0

US-10-718-311-8 (1-227) x US-10-450-763-26139 (1-1207)

QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76  
Db 634 TGTAAAGAGAGTTCGGCATGTACACCCCGGTTAAACCAACTGCGTGTGGCTAT 693  
QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96  
Db 694 TGTAAAGAGATCCCTCGCTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753  
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
Db 754 ATGACAAAACGTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGACGATGATCCCGAAGGG 813  
QY 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
Db 814 TTTGTGAGCAGAAATCAAAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 873  
QY 137 TrpLeuArgGluLeuLeu-LeuCysAlaAspGlyGluProTrpLeuAla-GlyArgThrV 156  
Db 874 TGGTTACGTAATTTTGTGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 933  
QY 156 alValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrP 176  
Db 934 TCGTTCTGTGTCAACGTTAAGCGGGCCGAGCTGGCGTTACAAAAATTTGGGTAAACGC 993

QY 176 roLeuGlyArgTyrIleuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyA 196  
Db 994 CGTTAGGACGCTATCTGTTCACATCATCGACATTAAACCGGAGCTTTATTGAGATAGGCC 1053  
QY 196 iqAspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuL 216  
Db 1054 GTGATGCCGGGCTGTGGGGGCGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTTC 1113  
QY 216 euThrGluLeuPheLeuProAlaSerProLeuTyr 227  
Db 1114 TAACAGAACTGTTTTTACCGGCGTCACCGTGTGAC 1148

Search completed: July 11, 2006, 22:12:04  
Job time : 1658.55 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 20:16:09 ; Search time 200.294 Seconds

(without alignments)  
3180.885 Million cell updates/sec

Title: US-10-718-311-8

Perfect score: 1159

Sequence: 1 MASSVISAANVATRNVTQA.....RLSGKPLLTFLFPASPLY 227

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10718311/runat 11072006 111527 5281/app query.fasta\_1  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs06p -USER=US10718311@CGN 1 1 381 @runat 11072006 111527 5281  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq:\*\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq:\*\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1159	100.0	684	3	US-09-896-866B-16
2	1159	100.0	684	3	US-09-855-341-7
3	856	73.9	854	3	US-09-064-693A-24
4	856	73.9	6641	3	US-09-064-693A-25
5	854	73.7	495	3	US-09-896-866B-12
6	854	73.7	495	3	US-09-855-341-3
7	673	58.1	645	3	US-09-489-039A-1234
8	401	34.6	537	3	US-09-543-681A-3122

9	277.5	23.9	499	3	US-09-495-797-41	Sequence 41, Appl
10	269	23.2	204	3	US-09-839-477-3	Sequence 3, Appl
11	262.5	22.6	1442	2	US-08-152-483B-8	Sequence 8, Appl
12	254	21.9	177	2	US-08-095-726-12	Sequence 12, Appl
13	254	21.9	177	2	US-08-096-043-12	Sequence 12, Appl
14	254	21.9	177	2	US-08-093-577-8	Sequence 8, Appl
15	254	21.9	177	2	US-08-331-004A-6	Sequence 6, Appl
16	254	21.9	177	2	US-08-096-623A-12	Sequence 12, Appl
17	254	21.9	177	7	PCT-US95-13937A-6	Sequence 6, Appl
18	210	18.1	10339	3	US-09-186-002-13	Sequence 13, Appl
19	204	17.6	10249	3	US-09-186-002-14	Sequence 14, Appl
20	203.5	17.6	696	3	US-09-441-340-11	Sequence 11, Appl
21	203.5	17.6	1630	3	US-09-441-340-23	Sequence 23, Appl
22	203	17.5	1178	3	US-09-434-039A-36	Sequence 36, Appl
23	203	17.5	1224	3	US-09-434-039A-34	Sequence 34, Appl
24	203	17.5	8012	3	US-09-182-117-1	Sequence 1, Appl
25	203	17.5	8012	3	US-09-434-039A-1	Sequence 1, Appl
26	203	17.5	8418	3	US-09-182-117-5	Sequence 5, Appl
27	203	17.5	8418	3	US-09-434-039A-5	Sequence 5, Appl
28	203	17.5	8798	3	US-09-182-117-4	Sequence 4, Appl
29	203	17.5	8798	3	US-09-434-039A-4	Sequence 4, Appl
30	201	17.3	264	3	US-09-441-340-9	Sequence 9, Appl
31	201	17.3	268	3	US-09-186-002-5	Sequence 5, Appl
32	201	17.3	279	2	US-08-391-339-9	Sequence 9, Appl
33	201	17.3	279	2	US-08-484-274A-9	Sequence 9, Appl
34	201	17.3	279	9	US-09-612-404-9	Sequence 9, Appl
35	201	17.3	355	2	US-08-090-523-5	Sequence 5, Appl
36	201	17.3	355	2	US-08-398-627-5	Sequence 5, Appl
37	201	17.3	355	2	US-08-406-858-5	Sequence 5, Appl
38	201	17.3	355	2	US-08-476-519-7	Sequence 7, Appl
39	201	17.3	355	3	US-08-399-023-5	Sequence 5, Appl
40	201	17.3	355	3	US-08-120-703A-5	Sequence 5, Appl
41	201	17.3	355	7	PCT-US91-04036-5	Sequence 5, Appl
42	201	17.3	355	7	PCT-US94-05275-5	Sequence 5, Appl
43	201	17.3	355	7	PCT-US95-09323-7	Sequence 7, Appl
44	201	17.3	509	3	US-09-011-151-1	Sequence 1, Appl
45	201	17.3	10846	3	US-09-098-219B-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-896-866B-16  
; Sequence 16, Application US/09896866B

; Patent No. 6642036

; GENERAL INFORMATION:

; APPLICANT: Flint, Dennis

; APPLICANT: Meyer, Knut

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Sinapylglucose:Malate Sinapyltransferase Form Malate Conjugates

; FILE REFERENCE: BC1034 US NA

; CURRENT APPLICATION NUMBER: US/09/896,866B

; CURRENT FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16

; LENGTH: 684

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei

US-09-896-866B-16

Alignment Scores:	3.84e-133	Length:	684
Pred. No.:	1159.00	Matches:	227
Score:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0



```

; ADDRESSES: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: LIT-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
; TELEFAX: (208)526-8339
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-064-693A-24
Alignment Scores:
Pred. No.: 1,31e-95 Length: 854
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-064-693A-24 (1-854)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrClnLeuArgAlaLeuArgTyr 76
Db 303 TGTAAACGGAGAGTTCGGCATGTACACCCCGCGTTAACCGCAACTGCGTGGCTGCCGTAT 362
QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 363 TGTAAAGAGATCCCTGCCCTGGATCGCAACTGCTCGACTGCTGTGTGGAGGATTC 422
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetLeuArgGluGly 116
Db 423 ATGACAAAACGTTTTTGAACAGCAGGAGAAACCGTAAGCGTGACGATGATCCCGAAGGG 482
QY 117 PheValGluGlnAsnGluIleProGluLeuLeuProLeuLeuProLysGluSerArgTyr 136
Db 483 TTTGTCGACAGCAATGAATCCCGGAGAACTGCGCGCTGCTGCGAAGAGTCTCGTTAC 542
QY 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 543 TGGTTACGTGAATTTTGTATGTGCGGATGTGACCGTGGCTGCGCGGTGATCCGTC 602
QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 603 GTTCTGTGTCAACGTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACCGCG 662
QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196
Db 663 TTAGGACGCTATCTGTTCATCATCATCAATTAAACCGGAGCTTTATTGAGATAGCCGT 722
QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
;
Db 723 GATCGCGGCTGTGGGGCGACGTTCCCGCCTGCATTAAGCGGTAACCGCTGTTGCTA 782
QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 783 ACAGAACTGTTTTTACCGCGGTCAACCGTTGTAC 815
RESULT 4
US-09-064-693A-25
; Sequence 25, Application US/09064693A
; Patent No. 6210937
; GENERAL INFORMATION:
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Gary Goodson
; ADDRESSEE: INEL--Lockheed Martin Idaho
; ADDRESSEE: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: LIT-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
; TELEFAX: (208)526-8339
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-064-693A-25
Alignment Scores:
Pred. No.: 3,79e-94 Length: 6641
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-064-693A-25 (1-6641)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrClnLeuArgAlaLeuArgTyr 76
Db 303 TGTAAACGGAGAGTTCGGCATGTACACCCCGCGTTAACCGCAACTGCGTGGCTGCCGTAT 362
QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 363 TGTAAAGAGATCCCTGCCCTGGATCGCAACTGCTCGACTGCTGTGTGGAGGATTC 422
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetLeuArgGluGly 116
Db 423 ATGACAAAACGTTTTTGAACAGCAGGAGAAACCGTAAGCGTGACGATGATCCCGAAGGG 482
```

Qy 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
Db 483 TTITGCGAGCAGAAATGAAATCCCGAAGAACTGCCGCTGCTCCCGAAAGAGTCTCGTTAC 542  
Qy 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156  
Db 543 TGGTTACGTGAATTTTGTATGCGCATGTTGAACCGTGGCTTCCCGCTGATACCGTC 602  
Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176  
Db 603 GTTCCTGTGTCAACGTTAAGCGGGCGGAGCTGGGCTTACAAAAATTTGGTAAAAACGCCG 662  
Qy 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluLeuGlyArg 196  
Db 663 TTAGAGCGCTATCTGTTACATCATCGACATTAAACCGGGACTTTATTGAGATAGCGCGT 722  
Qy 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216  
Db 723 GATCGCGGCTGTGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTA 782  
Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227  
Db 783 ACAGAACTGTTTTTACCAGCGTCACCGTTGTAC 815  
RESULT 5  
US-09-896-866B-12  
; Sequence 12, Application US/09896866B  
; Patent No. 6642036  
; GENERAL INFORMATION:  
; APPLICANT: Flint, Dennis  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugated  
; FILE REFERENCE: Benzoic Acid Glucosides  
; CURRENT APPLICATION NUMBER: US/09/896,866B  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/216,615  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-896-866B-12  
Alignment Scores:  
Pred. No.: 9,49e-96 Length: 495  
Score: 854.00 Matches: 165  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 73.7% Indels: 0  
DB: 3 Gaps: 0  
US-10-718-311-8 (1-227) x US-09-896-866B-12 (1-495)  
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82  
Db 1 ATGTACACACCCCGCGTTAACGCACTGCGTGGCTGCTATTGTAAGAGATCCCTGCC 60  
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 120  
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGAGGGGAAAAACCGTGAAGCTGATGATCCCGAAGGGTTTGTGCGAGCAGAAATGA 180  
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142  
Db 181 ATCCCGGAAAGAAACCGTGAAGCTGATCCCGAAGGGTTTGTGCGAGCAGAAATGA 180  
Qy 143 LeuProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142  
Db 181 ATCCCGGAAAGAAACCGTGAAGCTGATCCCGAAGGGTTTGTGCGAGCAGAAATGA 240

Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162  
Db 241 TTATGTCCGATGGTGAACCGTGGCTTCCCGTTCCTGTTGTTCAACGTTA 300  
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
Db 301 AGCGGGCGGAGCTGGCGTTACAAAAATTTGGTAAAAACCGCTTAGGACGCTATCTGTTT 360  
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCGCGTGTGCGGGTGTGGGG 420  
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuGluLeuPheLeuPro 222  
Db 421 CGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTTGTCTACAGAACTGTTTTTACCG 480  
Qy 223 AlaSerProLeuTyr 227  
Db 481 GCGTCACCGTTGTAC 495  
RESULT 6  
US-09-855-341-3  
; Sequence 3, Application US/09855341  
; Patent No. 6683231  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 3  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:E. coli  
US-09-855-341-3  
Alignment Scores:  
Pred. No.: 9,49e-96 Length: 495  
Score: 854.00 Matches: 165  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 73.7% Indels: 0  
DB: 3 Gaps: 0  
US-10-718-311-8 (1-227) x US-09-855-341-3 (1-495)  
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82  
Db 1 ATGTACACACCCCGCGTTAACGCACTGCGTGGCTGCTATTGTAAGAGATCCCTGCC 60  
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 61 CTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 120  
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 121 CAGCAGGAAAAACCGTGAAGCTGATGATCCCGAAGGGTTTGTGCGAGCAGAAATGA 180  
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142  
Db 181 ATCCCGGAAAGAAACCGTGAAGCTGATCCCGAAGGGTTTGTGCGAGCAGAAATGA 240  
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162  
Db 241 TTATGTCCGATGGTGAACCGTGGCTTCCCGTTCCTGTTGTTCAACGTTA 300



QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
Db 301 AGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACCGCGTTAGGACGCTATCTGTTC 360  
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
Db 361 ACATCATCGACATTAAACCGGAGCTTTATTGAGTAGGCGGCGTATCCGGCGCTGTGGGG 420  
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
Db 421 CGAGCGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 480  
QY 223 AlaSerProLeuTyr 227  
Db 481 GCGTCACCGTTGTAC 495

## RESULT 7

US-09-489-039A-1234  
; Sequence 1234, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 1234  
; LENGTH: 645  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-1234

## Alignment Scores:

Pred. No.: 3,11e-73 Length: 645  
Score: 673.00 Matches: 129  
Percent Similarity: 87.9% Conservative: 16  
Best Local Similarity: 78.2% Mismatches: 20  
Query Match: 58.1% Indels: 0  
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-489-039A-1234 (1-645)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
Db 148 ATGTCCCATCCTCGCTTACGCGACTGCGTGCCTGCTATTTTCGGTATGCGCATCC 207  
QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
Db 208 CTGCGCGCGCGCTAGTACTGCTGCTGCTGAGGACTCAATGACCCAGCGCTTGAA 267  
QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
Db 268 CAACAAGGAAGCAGGTACCGTGCCTGGTAAACGAGGATATATCGCGCGCTGACGG 327  
QY 123 IleProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
Db 328 CTGACTGACGACGCGGCGCTGCTGCGCGAGCGCGCTACTGCTGCGGAGATTATA 387  
QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162  
Db 388 CTCATGCGCATGCGAGCGCTGCTGCGCGCGCGCACGGTGGCGCGAGTCGACGCTG 447  
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
Db 448 TGTGGTCCGAGCTGGCGCTACAGCAGCTCGGCAAACTCCGCTGGCGCGGATACCTGTTT 507  
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202

Db 508 ACCTCGTCGACATTAAACCGCGATTTATTGAAATTTGTCGCGATCGAGCGTGTGGGG 567  
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
Db 568 CGTGTCTCCCGCTCGGCTGAGCGCAAAACCCCTGCTGACCGAGCTTTTTTTCCT 627  
QY 223 AlaSerProLeuTyr 227  
Db 628 GCGTCGCGCTGTGTAC 642

## RESULT 8

US-09-543-681A-3122  
; Sequence 3122, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 3122  
; LENGTH: 537  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-3122

## Alignment Scores:

Pred. No.: 8.23e-40 Length: 537  
Score: 401.00 Matches: 81  
Percent Similarity: 68.6% Conservative: 15  
Best Local Similarity: 57.9% Mismatches: 44  
Query Match: 34.6% Indels: 0  
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-543-681A-3122 (1-537)

QY 88 LeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThr 107  
Db 109 TTAAGTTGCTGCTGGAATTAGTTCATGACAGCAGCTTTTGAGCAACATTGCCATCAA 168  
QY 108 ValSerValThrMetIleArgGluGlyPheValGluGlnAsnGluIleProGluGluLeu 127  
Db 169 GTGACGGTAATGCCCATATCAAGAAGGTTTTATTGAATATATCGAGCCTGCTGATGAACAA 228  
QY 128 ProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeuLeuCysAlaAspGly 147  
Db 229 AAGTGTTTACCTTATAGCGCAGCTATTGCTTAAGAGAATCTCTTTTGTGGGATAAT 288  
QY 148 GluProTrpLeuAlaGlyArgThrValProValSerThrLeuSerGlyProGluLeu 167  
Db 289 GTACCTTGGTTATTAGGCGCAACATTAGTCCCGAAGACACATTAAACGGGTGAAGATCGC 348  
QY 168 AlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187  
Db 349 CAATTAGTGAATTTTTCGTACGCGTCCATTAGGACGTTATCTGTTCAGGAAACAACCTTA 408  
QY 188 ThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGlyArgArgSerArgLeu 207  
Db 409 AGTCGTGATTTTATTTCATATTGGGCAACAAATGGACATTGTTAGCTGTTCCCGTTTC 468  
QY 208 ArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuProLysSerProLeuTyr 227  
Db 469 CAGCTTTCAGATAAACCTTTATTATTACTGAGGTGTTTTTGCCTGCATCACCTGTATAT 528

## RESULT 9

US-09-495-797-41  
; Sequence 41, Application US/09495797  
; Patent No. 6369296  
; GENERAL INFORMATION:

```
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Newburn 43,047
; CURRENT FILING DATE: 2000-02-01
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence
US-09-495-797-41

Alignment Scores:
Pred. No.: 1,25e-24 Length: 499
Score: 277.50 Matches: 71
Percent Similarity: 61.5% Conservative: 17
Best Local Similarity: 45.7% Mismatches: 30
Query Match: 23.9% Indels: 25
DB: 3 Gaps: 7

US-10-718-311-8 (1-227) x US-09-495-797-41 (1-499)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCCTCAGTCTCTTCCTCTGCAGCAGTTCGCCCGCAGCAATGTTGCTCAAGCT 67
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 AACATGGTGGACCTTCCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 127
QY 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTGACATCACTTCCATTGCCAGCAACCGCGGAGAGTGCATGCAATGCATGCAGGTG 187
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArg-----AlaLeuArgTyrCysLys 78
Db 188 TGG-----CCCAATTAAACAGAGAGATGACGACTCTCTCATACCTTCTCT 235
QY 79 GlnIleProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSer 96
Db 236 GATTGAGC-----CAGGAGCAATGCTTAGTGAAGTTGAGTACCTTTTGAARAATGA 289
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 290 TGGGTTCCTTGTGTAATTCGAG-----ACTGAGCAGCGA 325
QY 117 PheVal-----GluGlnAenGlnIleProGluGluLeuProLeuProLysGluSer 134
Db 326 TTGTGCTACCGTGAACAACCAAGTCAACCAAGTCACTAT-----GATGCG 370
QY 135 ArgTyrTrp 137
Db 371 AGATACTGG 379

RESULT 10
US-09-839-477-3
; Sequence 3, Application US/09839477
; Patent No. 6723895
; GENERAL INFORMATION:
; APPLICANT: Shorrosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: COA-CARBOXYLASE
; FILE REFERENCE: 07148-094001
; CURRENT FILING DATE: 2001-04-20
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-09-839-477-3

Alignment Scores:
Pred. No.: 3.23e-24 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 23.2% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-839-477-3 (1-204)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCAGTCTCTTCCTCTGCAGCAGTTCGCCCGCAGCAATGTTGCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGACCTTCCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 120
QY 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACCGCGGAGAGTGCATGCAATGCATGCAGGTG 180
QY 61 Trp 61
Db 181 TGG 183

RESULT 11
US-08-152-483B-8
; Sequence 8, Application US/08152483B
; Patent No. 5529909
; GENERAL INFORMATION:
; APPLICANT: Della-Cioppa, Guy
; APPLICANT: Kumagai, Monto
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
; TITLE OF INVENTION: PROTEIN FUSION ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,483B
; FILING DATE: No. 5529909ember 12, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7/857,602
; FILING DATE: March 30, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 923,692
; FILING DATE: July 31, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 600,244
; FILING DATE: October 22, 1990
```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 641,617  
;; FILING DATE: January 16, 1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 737,899  
;; FILING DATE: July 26, 1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Halluin, Albert P.  
;; REGISTRATION NUMBER: 25,227  
;; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 854-3660  
;; TELEFAX: (415) 854-3694  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1442  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; DESCRIPTION: DNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; IMMEDIATE SOURCE:  
;; CLONE:  
;; FEATURE:  
US-08-152-483B-8

Alignment Scores:  
Pred. No.: 5,05e-22 Length: 1442  
Score: 262.50 Matches: 56  
Percent Similarity: 85.1% Conservative: 7  
Best Local Similarity: 75.7% Mismatches: 4  
Query Match: 22.6% Indels: 7  
DB: 2 Gaps: 1

US-10-718-311-8 (1-227) x US-08-152-483B-8 (1-1442)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 9 ATGGCTTCTCAGTTCTTCTCTGAGCAGTGTCCACCGGAGCAATGTGCTCAAGCT 68  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 69 AACATGGTTGCACCTTCTACCTGGCTTAAGTCAGCTGCCTCATTCCTCTTTCAAGGAG 128  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
Db 129 CAAACCTTTGACATCACTTCCATTTGCAGCAACGGGGAGAGTGAATGCAATG----- 182  
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeu 74  
Db 183 -----CCGGAACCTCACCCGTCGTCGCGCTC 209

RESULT 12  
US-08-095-726-12  
; Sequence 12, Application US/08095726  
; Patent No. 5530188  
; GENERAL INFORMATION:  
; APPLICANT: Ausich, Rodney L  
; APPLICANT: Brinkhaus, Friedhelm L  
; APPLICANT: Mukharji, Indrani  
; APPLICANT: Proffitt, John H  
; APPLICANT: Yarger, James G  
; APPLICANT: Yen, Huel-Che B  
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in  
; TITLE OF INVENTION: Genetically Engineered Hosts  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept

;; STREET: 200 E Randolph St  
;; CITY: Chicago  
;; STATE: IL  
;; COUNTRY: USA  
;; ZIP: 60680-0703  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.24  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/095,726  
;; FILING DATE: 21-JUL-1993  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/785,566  
;; FILING DATE: 30-OCT-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Galloway, No. 5530188val B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 3128567180  
;; TELEFAX: 3128564972  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 177 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-08-095-726-12

Alignment Scores:  
Pred. No.: 1.81e-22 Length: 177  
Score: 254.00 Matches: 51  
Percent Similarity: 96.6% Conservative: 6  
Best Local Similarity: 86.4% Mismatches: 2  
Query Match: 21.9% Indels: 0  
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x US-08-095-726-12 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 1 ATGGCTTCTCAGTTCTTCTCTGAGCAGTGTCCACCGGAGCAATGTGCTCAAGCT 60  
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 61 AACATGGTTGCACCTTCTACCTGGCTTAAGTCAGCTGCCTCATTCCTCTTTCAAGGAG 120  
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59  
Db 121 CAAACCTTTGACATCACTTCCATTTGCAGCAACGGGGAGAGTGAATGCAATGCAATG 177

RESULT 13  
US-08-096-043-12  
; Sequence 12, Application US/08096043  
; Patent No. 5530189  
; GENERAL INFORMATION:  
; APPLICANT: Ausich, Rodney L  
; APPLICANT: Brinkhaus, Friedhelm L  
; APPLICANT: Mukharji, Indrani  
; APPLICANT: Proffitt, John H  
; APPLICANT: Yarger, James G  
; APPLICANT: Yen, Huel-Che B  
; TITLE OF INVENTION: Lycopene Biosynthesis in  
; TITLE OF INVENTION: Genetically Engineered Hosts  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept  
; STREET: 200 E Randolph St  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA

```

; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,043
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,568
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530189val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-096-043-12

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x US-08-096-043-12 (1-177)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTCTTCTCTGCAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuApIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCAATGCATGCAG 177

RESULT 14
US-08-093-577-8
; Sequence 8, Application US/08093577
; Patent No. 5545816
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huel-Che B
; TITLE OF INVENTION: Phytoene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,577
; FILING DATE: 19-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,569
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5545816val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-093-577-8

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x US-08-093-577-8 (1-177)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTCTTCTCTGCAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuApIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCAATGCATGCAG 177

RESULT 15
US-08-331-004A-6
; Sequence 6, Application US/08331004A
; Patent No. 5618988
; GENERAL INFORMATION:
; APPLICANT: Hauptmann, Randal
; APPLICANT: Eschenfeldt, William H
; APPLICANT: English, Jami
; APPLICANT: Brinkhaus, Friedhelm L
; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
; TITLE OF INVENTION: in Storage Organs of Genetically
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corporation, Law Dept
; STREET: 55 Shuman Boulevard, Suite 600
; CITY: Naperville
; STATE: IL
; COUNTRY: USA
; ZIP: 60563-8437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,004A
; FILING DATE:

```

```

; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5618988val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7087172447
; TELEFAX: 7087172430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-331-004A-6

```

```

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

```

US-10-718-311-8 (1-227) x US-08-331-004A-6 (1-177)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCAGTTCTTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCTTTCACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCACTTCCTATTCGACGACGACGACGACGACGACGACGACGACGAC 177

```

Search completed: July 11, 2006, 21:02:56  
Job time : 204.294 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 11, 2006, 19:48:48 ; Search time 5789.68 Seconds  
(without alignments)  
3288.703 Million cell updates/sec

Title: US-10-718-311-8  
Perfect score: 1159  
Sequence: 1 MASSVISAANVATRSNVTTQA.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-MODEL=frame+ p2n.model -US10718311/runat 11072006 111523 5213/app query.fasta\_1  
-Q=/abs/ABSSWEB spool/SUFFIX-p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -TRANS=human40.cdi -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -ALIGN=15 -MODE=LOCAL  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p  
-USER=US10718311@CCN\_1\_12067 @runat 11072006 111523 5213 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

EST:\*

1:	gb_est1.*
2:	gb_est3.*
3:	gb_est4.*
4:	gb_est5.*
5:	gb_est6.*
6:	gb_hic.*
7:	gb_est2.*
8:	gb_est7.*
9:	gb_est8.*
10:	gb_est9.*
11:	gb_gse1.*
12:	gb_gse2.*
13:	gb_gse3.*
14:	gb_gse4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	618	53.3	832	13	CL666350 PRI0152b
2	510	44.0	828	13	CL687382 PRI0146c
3	320	27.6	504	7	AW040748 EST283612
4	316	27.3	431	7	AW623440 EST321385

5	316	27.3	554	1	AI777142	AI777142	EST258107
6	314	27.1	439	7	AW041050	AW041050	EST233914
7	314	27.1	456	7	AW091579	AW091579	EST284855
8	314	27.1	612	7	AW622820	AW622820	EST306806
9	313.5	27.0	420	1	AW041064	AW041064	EST283928
10	313.5	27.0	422	1	AI775340	AI775340	EST256440
11	313	27.0	449	7	AW093612	AW093612	EST286792
12	313	27.0	533	7	AW622747	AW622747	EST306884
13	313	27.0	621	2	BG129744	BG129744	EST475390
14	312.5	27.0	416	7	AW094716	AW094716	EST287896
15	312	26.9	434	3	BP896321	BP896321	BP896321
16	312	26.9	436	1	AI773402	AI773402	EST254502
17	312	26.9	436	7	AW039480	AW039480	EST281761
18	312	26.9	437	7	AW096716	AW096716	EST289896
19	312	26.9	439	7	AW039685	AW039685	EST282133
20	312	26.9	441	7	AW220510	AW220510	EST297063
21	312	26.9	443	7	AW037399	AW037399	EST278901
22	312	26.9	446	1	AI773933	AI773933	EST255033
23	312	26.9	446	1	AI899617	AI899617	EST269060
24	312	26.9	446	7	AW038809	AW038809	EST280765
25	312	26.9	446	7	AW039508	AW039508	EST281789
26	312	26.9	446	7	AW093842	AW093842	EST287022
27	312	26.9	447	2	BI929385	BI929385	EST549274
28	312	26.9	447	7	AW615833	AW615833	EST325283
29	312	26.9	449	1	AI772486	AI772486	EST253586
30	312	26.9	449	7	AW039228	AW039228	EST281463
31	312	26.9	449	7	AW094548	AW094548	EST287728
32	312	26.9	449	7	AW399384	AW399384	EST309884
33	312	26.9	452	7	AW037501	AW037501	EST276722
34	312	26.9	452	7	AW039295	AW039295	EST281552
35	312	26.9	453	7	BE462707	BE462707	EST325036
36	312	26.9	454	1	AI484250	AI484250	EST248477
37	312	26.9	454	1	AI490880	AI490880	EST241589
38	312	26.9	454	7	AW037928	AW037928	EST279572
39	312	26.9	456	1	AI772090	AI772090	EST253190
40	312	26.9	456	1	AI779629	AI779629	EST260508
41	312	26.9	470	3	BP889601	BP889601	BP889601
42	312	26.9	475	3	BP898763	BP898763	BP898763
43	312	26.9	476	3	BP905268	BP905268	BP905268
44	312	26.9	476	3	BP910180	BP910180	BP910180
45	312	26.9	477	3	BP909810	BP909810	BP909810

## ALIGNMENTS

RESULT 1  
CL666350/c

LOCUS  
DEFINITION  
CL666350  
CL666350.1 GI:50159025  
GSS.  
Pristionchus pacificus  
Pristionchus pacificus  
Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
survey sequence.  
1 (bases 1 to 832)  
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
AppDB: an AcedB database for the nematode satellite organism  
Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
14681447  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spananstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.

```

This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES             source
    Location/Qualifiers
     1..832
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
Alignment Scores:
Pred. No.:      5.67e-57      Length:      832
Score:          618.0%      Matches:      120
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 53.3%      Mismatches: 0
Query Match:      13         Indels:      0
DB:              13         Gaps:       0
US-10-718-311-8 (1-227) x CL666350 (1-832)
QY 108 ValSerValThrMetIleArgGluGlyPheValGluGlnAsnGluIleProGluGluLeu 127
    |||||
DB 832 GTAACGGTGACGATGATCCGCGAAGGGTTTGTGAGAGAGATGAAATCCCGAAGACTG 773
QY 128 ProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeuLeuCysAlaAspGly 147
    |||||
DB 772 CGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTGTATGTGCCGATGGT 713
QY 148 GluProTtpLeuAlaGlyArgThrValProValSerThrLeuSerGlyProGluLeu 167
    |||||
DB 712 GNACCGTGGCTGCGCGTGTACCGTCTCTGTGTCAAGTTAAGCGGCGGAGGTG 653
QY 168 AlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187
    |||||
DB 652 GCGTTACAAAATTTGGTAAACCGCGTTAGGACGCTATCTGTTACATCATCGACATTA 593
QY 188 ThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGlyArgArgSerArgLeu 207
    |||||
DB 592 ACCCGGGACTTTATTAGATAGCGCGTGATGCCGGGCTGTGGGGCGGACGTTCCCGCGCTG 533
QY 208 ArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuProAlaSerProLeuTyr 227
    |||||
DB 532 CGATTAAAGCGGTAAACCGCTGTGTCTACAGAACTGTTTTTACCAGCGCTACCGTTGTAC 473

RESULT 2
LOCUS CL687382 828 bp DNA linear GSS 09-JUL-2004
DEFINITION PRI0146c_H03.2 - PRI0146c.BR (828) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL687382
VERSION CL687382.1 GI:50196135
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 828)
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism
JOURNAL Pristionchus pacificus
PUBMED Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT 14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES             source
    Location/Qualifiers
     1..828
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
Alignment Scores:
Pred. No.:      4.26e-45      Length:      828
Score:          510.00      Matches:      99
Percent Similarity: 95.2%      Conservative: 1
Best Local Similarity: 94.3%      Mismatches: 5
Query Match:      13         Indels:      0
DB:              13         Gaps:       0
US-10-718-311-8 (1-227) x CL687382 (1-828)
QY 57 CysMetGlnValTtpHisMetSerHisProAlaLeuThrGlnLeuAlaLeuArgTyr 76
    |||||
DB 514 TGTAAACGGAGAGTTCGGCATGTCAACCCCGGTTAAGCAACTGCGTGGCGCTAT 573
QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTtpLeuLeuGluAspSer 96
    |||||
DB 574 TTTAAAGAGATCTCGCTCGATCCGCAACTGTCGACTGGCTGCTTCTGAGGATTCC 633
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
    |||||
DB 634 ATGCAAAACGTTTGAACAGCAGCGGAAAAACGGTAAGCGTGAAGATATCCGCAAGGG 693
QY 117 PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
    |||||
DB 694 TTTGTGAGCAGATGAATATCCCGAGAACTGCGGTGCTGCCGAAAGAGTCTGTTAC 753
QY 137 TrpLeuArgGluIleLeuLeuLeuCysAlaAspGlyGluProTtpLeuAlaGlyArgThrVal 156
    |||||
DB 754 TGGTTACGTGAAATTTGTTATGTGCGCATGCTGAACCGTGGCTTCCGCGTACCGCTC 813
QY 157 ValProValSerThr 161
    |||||
DB 814 GTTCTGTGTCAACG 828

RESULT 3
LOCUS AW040748 504 bp mRNA linear EST 18-MAY-2001
DEFINITION EST283612 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET8G6, mRNA sequence.
ACCESSION AW040748
VERSION AW040748.1 GI:5899502
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 504)
AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE Generation of ESTs from tomato leaf tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

```



Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.  
**FEATURES**  
 Location/Qualifiers  
 1..504  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET8G6"  
 /issue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MFG"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

# ORIGIN

Alignment Scores:  
 Pred. No.: 1-59e-24 Length: 504  
 Score: 320.00 Matches: 84  
 Percent Similarity: 60.4% Conservatives: 12  
 Best Local Similarity: 52.8% Mismatches: 39  
 Query Match: 27.6% Indels: 24  
 DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW040748 (1-504)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 12 ATGGCTTCCTCTGCTATTTCTTCAGCAGCTGTTGCCACACGCGCAATGTTACCAAGCT 71  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 72 AGCATGGTTGCACCTTCCTACTGCTCAATCTTCAGCCACTTCCTCTGTTCAAGAAG 131  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 DB 132 CAAACCTTGACATCATTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 191  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 DB 192 TGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCTGATTG 245  
 QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 DB 246 TCT-----GACGAGCAATTTGCTAGTGAATTTAGTACCTTTTGAATAATGATGGGTT 299  
 QY 99 LysArgPheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118  
 DB 300 CTTGCTTGGATTTCAG-----ACTGAGCAGCGAATTTGTC 335  
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 DB 336 TACCGTGAGAACACCAAGTCCACGAGTACTAT-----GATGGAAGGTAC 380  
 QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152  
 DB 381 TGGACCATGTGAAGTTGGCTATGTTGGTGTCATGTGATGCAACCCCACTGTTGGCT 437

RESULT 4  
 AW623440  
 LOCUS  
 DEFINITION  
 EST321385 tomato flower buds 3-8 mm, Cornell University  
 Lycopersicon esculentum cDNA clone cTOB10124 5', mRNA sequence.  
 AW623440  
 ACCESSION  
 VERSION  
 AW623440.1 GI:7336467  
 KEYWORDS  
 EST.  
 SOURCE  
 Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM  
 Lycopersicon

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE AUTHORS

1 (bases 1 to 431)  
 van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E.,  
 Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
 Niemman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
 Tanksley,S.D.  
 Generation of ESTs from tomato flower tissue, 3-8 mm buds  
 Unpublished (1999)

## TITLE JOURNAL COMMENT

Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

## FEATURES

Location/Qualifiers  
 1..431  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOB10124"  
 /tissue\_type="flower"  
 /dev\_stage="3-8mm buds"  
 /clone\_lib="tomato flower buds 3-8 mm, Cornell University"  
 /note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."

## source

US-10-718-311-8 (1-227) x AW623440 (1-431)

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.47e-24 Length: 431  
 Score: 316.00 Matches: 84  
 Percent Similarity: 60.4% Conservatives: 12  
 Best Local Similarity: 52.8% Mismatches: 39  
 Query Match: 27.3% Indels: 24  
 DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW623440 (1-431)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

DB 6 ATGGCTTCCTCTGCTATTTCTTCAGCAGCTGTTGCCACACGCGCAATGTTACCAAGCT 65

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

DB 66 AGCATGGTTGCACCTTCCTACTGCTCAATCTTCAGCCACTTCCTCTGTTCAAGAAG 125

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

DB 126 CAAACCTTGACATCATTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 185

QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80

DB 186 TGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCTGATTG 239

QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98

DB 240 TCT-----GACGAGCAATTTGCTAGTGAATTTAGTACCTTTTGAATAATGATGGGTT 293

QY 99 LysArgPheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118

DB 294 CTTGCTTGGATTTCAG-----ACTGAGCAGCGAATTTGTC 329

QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136

DB 330 TACCGTGAGAACACCAAGTCCACGAGTACTAT-----GATGGAAGGTAC 374

QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152

```

Db      375 TGGACCATGTGGAGTTCCTATGTTAGGGTGCACTGATGCAACCAAGTGTGGCT 431
RESULT 5
AI777142
LOCUS   EST258107 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION
AI777142
ACCESSION
AI777142
VERSION
AI777142.1 GI:5275183
KEYWORDS
EST.
SOURCE  Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 554)
AUTHORS
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
TITLE   Generation of ESTs from Pseudomonas resistant tomato
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..554
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER20020"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Alignment Scores:
Pred. No.: 5,07e-24 Length: 554
Score: 316.00 Matches: 84
Percent Similarity: 60.4% Conservative: 12
Best Local Similarity: 52.8% Mismatches: 39
Query Match: 27.3% Indels: 24
DB: 1 Gaps: 7

US-10-718-311-8 (1-227) x AI777142 (1-554)
Qy      1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db      1 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTGGCCACACGACGAAATGTTACACAAGCT 60

Qy      21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrIleLys 40
Db      61 AGCATGGTTGGACCTTTCCACTGGTCTCAAAATCTTCAGCCACCTTTCCCTGTTTACAAAGAAG 120

Qy      41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db      121 CAAACCTTGACATCATCTTCATTGCTAGCATGTGGTGAAGAGTTAGCTGATGACAGGTG 180

Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      181 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCCTCGATTG 234

Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98

```

---

```

Db      235 TCT-----GACGAGCAATTGCTTAGTGAAATAGTACCTTTTGAATAATGGATGGGT 288
Qy      99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
Db      289 CCTTGCTTGGAAATTTAG-----ACTGACACGAGATTGTC 324
Qy      119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db      325 TACCGTGAGAACAAAGTCCACAGGATCTAT-----GATGAAGGTAC 369
Qy      137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
Db      370 TGGACCATGTGGAAAGTTCCTATGTTGGGTGCACGTGATGCAACCAAGTGTGGCT 426

RESULT 6
AW041050
LOCUS   EST283914 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
DEFINITION
clone cLET11E11, mRNA sequence.
ACCESSION
AW041050
VERSION
AW041050.1 GI:5899804
KEYWORDS
EST.
SOURCE  Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 439)
AUTHORS
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE   Generation of ESTs from tomato leaf tissue
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..439
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone="cLET11E11"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="X11-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisocotinic
acid, BTH, jasmonic acid, ethylene, fenchion, EIX,
okadaic acid, or systemin prior to tissue harvest.
site was destroyed during cloning."

ORIGIN
Alignment Scores:
Pred. No.: 5.92e-24 Length: 439
Score: 314.00 Matches: 83
Percent Similarity: 59.7% Conservative: 12
Best Local Similarity: 52.2% Mismatches: 40
Query Match: 27.1% Indels: 24
DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW041050 (1-439)
Qy      1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db      9 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTGGCCACACGACGAAATGTTACACAAGCT 68

```

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 69 AGCATGGTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTCCCTGTACAAAGAAG 128  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyValSerCysMetGlnVal 60  
 |||||  
 Db 129 CAAACCTTGACATCACTTCCATCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGTG 188  
 |||||  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 |||||  
 Db 189 TGGCCACCAATTAAC-----ATGAAGAGTACGAGACACTCTCATACCTTCCTGATTG 242  
 |||||  
 QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 |||||  
 Db 243 TCT-----GACGAGCAATTCCTAGTAAATGAGTACCTTTTGAAAAATGATGGGTT 296  
 |||||  
 QY 99 LysArgPheGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118  
 |||||  
 Db 297 CCTTGGTGGATTTGAG-----ACTGAGCAGCGATTGTC 332  
 |||||  
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 |||||  
 Db 333 TACCGTGAGAACCAACGATCACCAGATACATAT-----GATGGAAGGTAC 377  
 |||||  
 QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152  
 |||||  
 Db 378 TGGACCATGGGAAGTTGGCTATGTTGGTGCGATGATGCAACCAAGGTGGCT 434  
 |||||

## RESULT 7

LOCUS AW091579  
 DEFINITION EST284855 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET15H6, mRNA sequence.  
 ACCESSION AW091579  
 VERSION 1  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

source

1. .456  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET15H6"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, ETX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

## ORIGIN

source

## Alignment Scores:

Pred. No.: 6,27e-24 Length: 456  
 Score: 314.00 Matches: 83  
 Percent Similarity: 59.7% Conservative: 12  
 Best Local Similarity: 52.2% Mismatches: 40  
 Query Match: 27.1% Indels: 24  
 DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW091579 (1-456)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 16 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 75  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 76 AGCATGGTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTCCCTGTACAAAGAAG 135  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyValSerCysMetGlnVal 60  
 |||||  
 Db 136 CAAACCTTGACATCACTTCCATTCCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGTG 195  
 |||||  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 |||||  
 Db 196 TGGCCACCAATTAAC-----ATGAAGAGTACGAGACACTCTCATACCTTCCTGATTG 249  
 |||||  
 QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 |||||  
 Db 250 TCT-----GACGAGCAATTCCTAGTAAATGAGTACCTTTTGAAAAATGATGGGTT 303  
 |||||  
 QY 99 LysArgPheGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118  
 |||||  
 Db 304 CCTTGGTGGATTTGAG-----ACTGAGCAGCGATTGTC 339  
 |||||  
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 |||||  
 Db 340 TACCGTGAGAACCAACGATCACCAGATACATAT-----GATGGAAGGTAC 384  
 |||||  
 QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152  
 |||||  
 Db 385 TGGACCATGGGAAGTTGGCTATGTTGGTGCGATGATGCAACCAAGGTGGCT 441  
 |||||

## RESULT 8

LOCUS AW622820  
 DEFINITION EST306806 tomato flower buds 3-8 mm, Cornell University Lycopersicon esculentum cDNA clone cTOB1M21 5', mRNA sequence.  
 ACCESSION AW622820  
 VERSION 1  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 612)

van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

Generation of ESTs from tomato flower tissue, 3-8 mm buds

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

Location/Qualifiers

1. .612

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

```

/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT081W21"
/tissue_type="flower"
/dev_stages="3-8mm buds"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tankale; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 9,77e-24 Length: 612
Score: 314.00 Matches: 83
Percent Similarity: 59.7% Conservative: 12
Best Local Similarity: 52.2% Mismatches: 40
Query Match: 27.1% Indels: 24
DB: 7 Gaps: 7

```

US-10-718-311-8 (1-227) x AW622820 (1-612)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 9 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGT 68

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 69 AGCATGGTGTGCACCTTCCTGCTCAATCTTCAGCCACTTTCCTCCCTTTACAAAGAAG 128

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 129 CAAAACCTTGACATCACTTCCATTCTAGCAATGCTGGAAGAGTTAGCTGCATGCAGTG 188

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeu 80
Db 189 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATCTTCTCGATTG 242

Qy 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db 243 TCT-----GACGACCAATTTGCTAGTAAATGAGTACCTTTTGAAAAAATGGATGGTT 296

Qy 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
Db 297 CCTTGCTTGGAAATTGAG-----ACTGAGCAGCGATTGCTC 332

Qy 119 -----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 333 TACCGTGAGAACAAACAGTCACCGATCACTAT-----GATGGAAGGTAC 377

Qy 137 TrpLeu-----ArgGluLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
Db 378 TGGACCATGTGGAAGTCGCCTATGTTGGGTGCATGTGCAACCAAGTGTGGCT 434

```

## RESULT 9

## AW041064

## LOCUS

EST281928 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

clone cLET11A11, mRNA sequence.

## ACCESSION

## AW041064

## VERSION

## AW041064.1

## KEYWORDS

## SOURCE

## ORGANISM

## Lycopersicon esculentum (Solanum lycopersicum)

## Eukaryota; Viridiplantae;

## Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

## asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## 1 (bases 1 to 420)

## D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,

## Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,

## Frazer,C.M., Venter,J.C., Martin,G.B., Tankaley,S.D. and

## TITLE

## JOURNAL

## COMMENT

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

## source

1..420

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/cultivar="Rio Grande PtoR"

/db\_xref="taxon:4081"

/clone="cLET11A11"

/tissue\_type="leaf"

/dev\_stages="4-6 week old plants"

/lab\_hosts="XLI-Blue MRF"

/clone\_lib="tomato mixed elicitor, BTI"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. site was destroyed during cloning."

## ORIGIN

## Alignment Scores:

```

Pred. No.: 6,28e-24 Length: 420
Score: 313.50 Matches: 79
Percent Similarity: 63.8% Conservative: 11
Best Local Similarity: 56.0% Mismatches: 30
Query Match: 27.0% Indels: 21
DB: 7 Gaps: 6

```

US-10-718-311-8 (1-227) x AW041064 (1-420)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 19 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGT 78

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 79 AGCATGGTGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCCCTTTACAAAGAAG 138

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 139 CAAAACCTTGACATCACTTCCATTGCTAGCAATGCTGGAAGAGTTAGCTGCATGCAGTG 198

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeu 80
Db 199 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATCTTCTCGATTG 252

Qy 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db 253 TCT-----GACGACCAATTTGCTAGTAAATGAGTACCTTTTGAAAAAATGGATGGTT 306

Qy 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
Db 307 CCTTGCTTGGAAATTGAG-----ACTGAGCAGCGATTGCTC 342

Qy 119 -----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 343 TACCGTGAGAACAAACAGTCACCAAGATCACTAT-----GATGGAAGGTAC 387

Qy 137 Trp 137
Db 388 TGG 390

```

## RESULT 10

## AI775340

## LOCUS

## DEFINITION

## EST256440 tomato resistant, Cornell Lycopersicon esculentum cDNA

```

clone cLER15A14, mRNA sequence.
AI775340
AI775340.1 GI:5273381
EST.
Lycopersicon esculentum (Solanum lycopersicum)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 422)
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .422
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Ril-12 (358::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER15A14"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."
ORIGIN
source
Alignment Scores:
Pred. No.: 6.33e-24 Length: 422
Score: 313.50 Matches: 79
Percent Similarity: 63.8% Conservative: 11
Best Local Similarity: 56.0% Mismatches: 30
Query Match: 27.0% Indels: 21
DB: 1 Gaps: 6
US-10-718-311-8 (1-227) x AI775340 (1-422)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 6 ATGGCTTCCTCTGTCATTTCTTTCCAGCAGCTGTGGCCACAGCAGCAATGTTACACAAAGCT 65
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 66 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAG 125
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 126 CAAACCTTTGACATCACTTCCATTTGTCAGCAATGGTGGAAAGAGTTAGTGCATGCAGGTG 185
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB 186 TGCCACCACCAATTAAC-----ATGAAGAAGTAGCAGACACTTCATPACCTTCCTGATTTG 239
QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
DB 240 TCT-----GACGAGCAATTCCTTAGTGAAATGAGTACCTTCTGAAAATGATGGGTT 293
QY 99 LysArgPheGlnGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
DB 294 CCTTGCTTGAATTTGAG-----ACTGAGCAGCAGATTTGTC 329
clone cLER15A14, mRNA sequence.
AI775340
AI775340.1 GI:5273381
EST.
Lycopersicon esculentum (Solanum lycopersicum)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 449)
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato leaf tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .449
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone="cLET25B3"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XL1-Blue MRF,"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
ORIGIN
Alignment Scores:
Pred. No.: 7.89e-24 Length: 449
Score: 313.00 Matches: 84
Percent Similarity: 59.7% Conservative: 11
Best Local Similarity: 52.8% Mismatches: 40
Query Match: 27.0% Indels: 24
DB: 7 Gaps: 7
US-10-718-311-8 (1-227) x AW093612 (1-449)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 9 ATGGCTTCCTCTGTCATTTCTTTCCAGCAGCTGTGGCCACAGCAGCAATGTTACACAAAGCT 68
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 69 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAG 128
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

```

```

Db      129 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 188
Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      189 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACACTCTCATACATCTCCCTGATTG 242
Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db      243 TCT-----GAGCAGCAATTCCTAGTGAATAGTACCTTTTGAAAAATGGATGGGTT 296
Qy      99 LysArgPheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
Db      297 CCTTGCTTGGAAATTTGAG-----ACTGACGACGGATTTGTC 332
Qy      119 -----GluGlnAenGluIleProGluGluLeuProLeuProLysGluSerArgTyr 136
Db      333 TACCGTGAGAACAAAGTCACCAAGGATACAT-----GATGGAAGGTAC 377
Qy      137 TrpLeuArgGluIleLeuLeu-----CysAlaAspGlyGluProTrpLeuAla 152
Db      378 TGGACCATGNTGAAGTTGCCTATGTTTGGGTGCATGATGCAACCAAGTGTGGCT 434

RESULT 12
AW622747 533 bp mRNA linear EST 18-MAY-2001
LOCUS EST306884 tomato flower buds 3-8 mm, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cTOB402 5', mRNA sequence.
ACCESSION AW622747
VERSION AW622747.1 GI:7335841
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 533)
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Frazer,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 3-8 mm buds
Unpublished (1999)
Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
1..533
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="cTOB402"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksey; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Alignment Scores:
Pred. No.: 1.02e-23 Length: 533
Score: 313.00 Matches: 83
Percent Similarity: 60.4% Conservative: 13
Best Local Similarity: 52.2% Mismatches: 39
Query Match: 27.0% Indels: 24
Db: 7 Gaps: 7

```

```

US-10-718-311-8 (1-227) x AW622747 (1-533)

Qy      1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla 20
Db      17 ATGCTTCCTCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 76
Qy      21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db      77 AGCATGGTTGCACCTTTCACCTGGTCTCAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 136
Qy      41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db      137 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 196
Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      197 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACACTCTCATACCTTCTCGATTG 250
Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db      251 TCT-----GAGCAGCAATTCCTTAATGAAATTCGATACCTTTTGAAAAATGGATGGGTT 304
Qy      99 LysArgPheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
Db      305 CCTTGCTTGGAAATTTGAG-----ACTGACGACGGATTTGTC 340
Qy      119 -----GluGlnAenGluIleProGluGluLeuProLeuProLysGluSerArgTyr 136
Db      341 TACCGTGAGAACAAAGTCACCAAGGATACAT-----GATGGAAGGTAC 385
Qy      137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
Db      386 TGGACCATGTTGAAGTTGCCTATGTTTGGGTGCATGATGCAACCAAGTGTGGCT 442

RESULT 13
BG129744 621 bp mRNA linear EST 31-JAN-2001
LOCUS EST475390 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION cTOF28E15 5' sequence, mRNA sequence.
ACCESSION BG129744
VERSION BG129744.1 GI:12629932
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 621)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..621
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="cTOF28E15"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

FEATURES
source
1..621
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clones="cTOF28E15"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

ORIGIN

```

## Alignment Scores:

Pred. No.: 1,29e-23 Length: 621  
 Score: 313.00 Matches: 94  
 Percent Similarity: 50.0% Conservative: 15  
 Best Local Similarity: 43.1% Mismatches: 61  
 Query Match: 27.0% Indels: 48  
 DB: 2 Gaps: 9

US-10-718-311-8 (1-227) x BG129744 (1-621)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 15 ATGGCTTCCTCTGTCTCAATTTCTTCAGCAGCTGTTCACAGCAAGCAATGTTACCAAGCT 74  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 75 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 134  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 135 CAAACCTTGACATCACTTCCATTTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 194  
 |||||  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 |||||  
 Db 195 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCCTGATTG 248  
 |||||  
 QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 |||||  
 Db 249 TCT-----GACGAGCAATGGTCTAGTGAATTTAGTACTTTTGAATAATGATGGTT 302  
 |||||  
 QY 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118  
 |||||  
 Db 303 CCTTGTCTGGAATTCAG-----ACTGAGCAGCGAATTGTC 338  
 |||||  
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 |||||  
 Db 339 TACCGTGTGAGAACCAAGTACCACGAGTACTAT-----GATGGAAGGTAC 383  
 |||||  
 QY 137 TrpLeu-----138  
 |||||  
 Db 384 TGGACCATGTGGGAAGTTCCTATGTTGGTGCACCTGATGCAACCAAGTGTGCTGA 443  
 |||||  
 QY 139 -----ArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeu-----151  
 |||||  
 Db 444 GGTTCAGAGCGCTAAAGGACATACCCACAGCATGGTGCAGAAATCATGATTGCACAA 503  
 |||||  
 QY 152 ---AlaGlyArgThrValProValSerThrLeuSerGlyProGluLeuAlaLeuGln 170  
 |||||  
 Db 504 TGTGGCTCAAGTGCAGTGTATCATGTTTCATGCTACAGCCAGAGCGCTACTAAGTTTC 563  
 |||||  
 QY 171 LysLeuGly---LysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187  
 |||||  
 Db 564 ATATTAGGAACAAATGCTTTAGGACACTTTGTTTAAATGCTACTTTA 617  
 |||||

## RESULT 14

AW094716 416 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST287896 tomato mixed elicitor, Bti Lycopersicon esculentum cDNA  
 DEFINITION

clone cLET29N8, mRNA sequence.

## ACCESSION

AW094716

## VERSION

1

## KEYWORDS

EST.

## SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 416)

## REFERENCE

D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,

Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W.,

Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and

Giovannoni, J.

## TITLE Generation of ESTs from tomato leaf tissue

## JOURNAL

## COMMENT

Unpublished (1999)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

## FEATURES

## source

1..416  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET29N8"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XL1-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, Bti"  
 /note="Vector: pBlueScript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7.98e-24 Length: 416  
 Score: 312.50 Matches: 78  
 Percent Similarity: 65.5% Conservative: 13  
 Best Local Similarity: 56.1% Mismatches: 31  
 Query Match: 27.0% Indels: 17  
 DB: 7 Gaps: 5

US-10-718-311-8 (1-227) x AW094716 (1-416)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 |||||  
 Db 6 ATGGCTTCCTCTGTCTCAATTTCTTCAGCAGCTGTTCACAGCAAGCAATGTTACCAAGCT 65  
 |||||  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 |||||  
 Db 66 ACATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 125  
 |||||  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 |||||  
 Db 126 CAAACCTTGACATCACTTCCATTTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 185  
 |||||  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80  
 |||||  
 Db 186 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCCTGATTG 239  
 |||||  
 QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98  
 |||||  
 Db 240 TCT-----GACGAGCAATTTGCTTAGTGAATTTAGTACTTTTGAATAATGATGGTT 293  
 |||||  
 QY 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118  
 |||||  
 Db 294 CCTTGTCTGGAATTTGAGACTGACGACGAGTCTGTC-----TACCGT 335  
 |||||  
 QY 119 GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrp 137  
 |||||  
 Db 336 GAGAACCAAGTACCCAGATACTAT-----GATGAAGGTACTGG 377  
 |||||

## RESULT 15

## BP896321

## LOCUS

## DEFINITION

clone LA08CH03 5', mRNA sequence.

## ACCESSION

BP896321

## VERSION

BP896321.1

## KEYWORDS

EST.







GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: July 11, 2006, 19:35:08 ; Search time 631.516 Seconds  
(without alignments)  
3759.291 Million cell updates/sec

Title: US-10-718-311-8  
Perfect score: 1159  
Sequence: 1 MASSVISNAVATRNVTQA.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10718311/runat\_11072006\_111518\_5144/app\_query.fasta.1  
-DB=N Geneseq -OFT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss05p  
-USER=US10718311@CGN\_1\_1423 @runat\_11072006\_111518\_5144 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_8.\*

```
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*
15: Geneseqn2006s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1159	100.0	684	6	ABA91838 Chloropla
2	1159	100.0	684	6	Abk47712 DNA seque
3	1159	100.0	684	10	Adc25988 Tomato/Es

ID	ABA91838	standard; DNA; 684 BP.	684	14	ADV91622	Tomato ru
4	1159	100.0	684	14	ADV91622	Adv91622 Tomato ru
5	862	74.4	3452	14	ADV91636	Adv91636 C. paraps
6	856	73.9	854	4	AAD06956	Aad06956 ubiC gene
7	856	73.9	6641	4	AAD06957	Aad06957 pME2 comp
8	854	73.7	495	2	AAQ92409	Aaq92409 E.coli ch
9	854	73.7	495	6	ABA91837	Abk47709 DNA seque
10	854	73.7	495	6	ABK47709	Abk47709 DNA seque
11	854	73.7	495	10	ADC25984	Adc25984 Escherich
12	854	73.7	498	14	ADV91618	Adv91618 Escherich
13	853	73.6	2000	2	AAT29820	Aat29820 E. coli u
14	851	73.4	1971	14	ADV91643	Adv91643 C. paraps
15	813	70.1	1207	5	AAS90335	Aas90335 DNA encod
16	795	68.6	1218	2	AAQ67671	Aaq67671 E.coli ub
17	673	58.1	645	11	ACH95439	Ach95439 Klebsiell
18	417	36.0	519	10	ACF70957	Acf70957 Photorhab
19	417	36.0	110000	10	ACF67367	Acf67367 40
20	417	36.0	110000	10	ACF65388	Acf65388 07
21	401	34.6	537	10	ADF02837	Adf02837 Bacterial
22	282.5	24.4	615	10	ADC76949	Adc76949 DNA homol
23	281.5	24.3	608	10	ADC76944	Adc76944 DNA homol
24	281.5	24.3	684	10	ADC76948	Adc76948 DNA homol
25	281.5	24.3	859	11	ADM45065	Adm45065 Insect re
26	277.5	23.9	492	10	ADC76164	Adc76164 DNA homol
27	277.5	23.9	499	6	AD35635	Ad35635 Nicotiana
28	277.5	23.9	504	10	ADK59824	Adk59824 Plant DNA
29	277.5	23.9	619	11	ADM45158	Adm45158 Insect re
30	276.5	23.9	489	10	ADK56120	Adk56120 Plant DNA
31	276.5	23.9	581	10	ADC75089	Adc75089 N bentham
32	276.5	23.9	609	10	ADC76965	Adc76965 DNA homol
33	276.5	23.9	614	10	ADC76956	Adc76956 DNA homol
34	276.5	23.9	632	10	ADC76953	Adc76953 DNA homol
35	276.5	23.9	718	10	ADK58382	Adk58382 Plant DNA
36	276.5	23.9	736	10	ADK54321	Adk54321 Plant DNA
37	276.5	23.9	736	10	ADK57660	Adk57660 Plant DNA
38	276.5	23.9	841	11	ADM44879	Adm44879 Insect re
39	276.5	23.9	847	11	ADM45447	Adm45447 Insect re
40	269	23.2	204	6	AAI70688	Aai70688 Tobacco R
41	269	23.2	297	10	ADK59826	Adk59826 Plant DNA
42	269	23.2	377	10	ADK56121	Adk56121 Plant DNA
43	269	23.2	924	10	ADC75566	Adc75566 DNA homol
44	269	23.2	958	10	ADC76165	Adc76165 DNA homol
45	266	23.0	668	10	ADK59825	Adk59825 Plant DNA

#### ALIGNMENTS

RESULT 1  
ABA91838  
ID ABA91838 standard; DNA; 684 BP.  
XX ABA91838;  
AC ABA91838;  
XX 29-AUG-2003 (revised)  
DT 15-MAY-2002 (first entry)  
XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.  
XX Chloroplast transit peptide; Rubisco; plant;  
KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;  
XX enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.  
OS Lycopersicon esculentum.  
OS Escherichia coli.  
OS Chimeric.  
XX Key Location/Qualifiers  
FT CDS 1..495  
FT /tag= a  
FT /product= "chloroplast-targeted CPL fusion"  
FT transit\_peptide 1..186  
FT /tag= b  
FT /note= "tomato Rubisco chloroplast targeting sequence"  
FT mat\_peptide 187..492

```

FT      /*tag= c
XX      /note= "E. coli chorismate pyruvate lyase"
PN      WO200194607-A2.
XX      13-DEC-2001.
XX
XX      22-MAY-2001; 2001WO-US016661.
XX
XX      02-JUN-2000; 2000US-0209854P.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX      Meyer K, Van Dyk DE, Viitanen PV;
XX
XX      WPI; 2002-226795/28.
XX      P-PSDB; AAM50959.
XX
XX      Producing para-hydroxy benzoic acid in green plant, comprises expression
XX      of unique expression cassette containing gene encoding chorismate
XX      pyruvate lyase operably linked to specific chloroplast targeting
XX      sequence.
XX
XX      Claim 17; Page 57; 60pp; English.
XX
XX      The present nucleotide sequence is that of an open reading frame encoding
XX      chloroplast-targeted chorismate pyruvate lyase (CPL) fusion protein TP-
XX      CPL (see AAM50959). PCR (see ABA91841-42) was used to generate a DNA
XX      fragment corresponding to the transit peptide of tomato Rubisco small
XX      subunit and the first 4 amino acid residues of mature Rubisco. The DNA
XX      fragment was ligated into pET24a-CPL, which carries the Escherichia coli
XX      CPL open reading frame (see ABA91837). The TP-CPL construct is an example
XX      of expression cassettes of the invention that are designed for the high-
XX      level production of p-hydroxybenzoic acid (pHBA) in higher plants.
XX      Cleavage of TP-CPL in the chloroplast releases a novel polypeptide (see
XX      AAM50961) that has full enzyme activity, converting chorismate to pHBA. A
XX      claimed plant comprising a CPL expression cassette is selected from
XX      soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
XX      barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,
XX      millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
XX      component of liquid crystal polymers which have application in the
XX      automotive, electrical and other industries. (Updated on 29-AUG-2003 to
XX      standardize OS field)
XX
XX      SQ      Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      5,77e-119      Length:      684
Score:      1159.00      Matches:      227
Percent Similarity:      100.0%      Conservative:      0
Best Local Similarity:      100.0%      Mismatches:      0
Query Match:      100.0%      Indels:      0
DB:      6      Gaps:      0

US-10-718-311-8 (1-227) x ABA91838 (1-684)

QY      1      MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB      1      ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTTGCCACACGACGACGATGTTACACAGCT 60
QY      21      SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40
DB      61      AGCATGGTGGACCTTTCACTGGTCTCAAACTCTTCAGCCACTTTCCCTGTTTACAAGAAG 120
QY      41      GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetClnVal 60
DB      121      CAAACCTTGATCATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATCGAGGTG 180
QY      61      TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB      181      TGGCATATGTACACCCCGCGTTAAACGCAACTGCGTGGCTGCGCTATTGTAAGAGATC 240
QY      81      ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrIlyArg 100

```

---

```

Db      241      CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAGCT 300
QY      101      PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db      301      TTTGAACAGCAGGAGGAAACGGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTTCGAGCAG 360
QY      121      AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db      361      AATGAAATCCCCGAGAACTGCCGCTGCTCCGAAAGAGTCTCGTTACTGTTAGTGA 420
QY      141      IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
Db      421      ATTTTGTATGTGCGGATGGTGAACCGTGGCTTCCCGTTCGTACCTGTTCTCTGTGTC 480
QY      161      ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db      481      ACGTTAAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACGCCCTTAGGACGCTAT 540
QY      181      LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db      541      CTGTTTCATCATCATCGACATTAAACCCGGACTTTTATTGAGATAGCCCGGTGATGCCGGCTG 600
QY      201      TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db      601      TGGGCGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 660
QY      221      LeuProAlaSerProLeuTyr 227
Db      661      TTACCGGCGTCACCGTTGTATC 681

RESULT 2
ABK47712
ID      ABRK47712 standard; DNA; 684 BP.
XX
XX      ABK47712;
XX
XX      18-JUN-2002 (first entry)
XX
XX      DNA sequence of ORF for TP-CPL fusion protein.
XX
XX      SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester; plant;
XX      sinapoylglucose:maltate sinapoyltransferase; glycosylated aromatic acid;
XX      maltate conjugated aromatic acid; polymer synthesis; ubiC; TP-CPL; gene;
XX      carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;
XX      tomato Rubisco small subunit precursor; rbcS2; mutant; ds.
XX
XX      Escherichia coli.
XX      Lycopersicon esculentum.
XX      Synthetic.
XX      Chimeric.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..684
XX              /*tag= a
XX              /product= "TP-CPL fusion protein"

WO200204653-A2.
17-JAN-2002.
05-JUL-2001; 2001WO-US021283.
07-JUL-2000; 2000US-0216615P.
(DUPO ) DU PONT DE NEMOURS & CO E I.
Flint D, Meyer K, Viitanen PV;
WPI; 2002-303779/34.
P-PSDB; AAU77942.
Producing aromatic acid conjugates, involves contacting glycosylated

```

PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the  
 PT presence of sinapoylglucose:malate sinapoyltransferase.

PS Example; Page 70; 72pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana  
 CC gene (SNG1, sinapoylglucose accumulator 1) encoding  
 CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the  
 CC substitution of a glucose moiety on a glycosylated aromatic acid with a  
 CC malate moiety to form a malate conjugated aromatic acid. The enzyme is  
 CC useful for producing malate conjugated aromatic acids, carboxylic acid  
 CC conjugated aromatic acids or aromatic esters. The malate conjugated  
 CC aromatic acids are useful in the synthesis of various polymers. The  
 CC present sequence representing the ORF (open reading frame) for TP-CPL  
 CC (tomato Rubisco small subunit precursor for rbcS2-Escherichia coli  
 CC chorismate pyruvate lyase (CPL)) fusion protein is described in the  
 CC examples of the present invention

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 5,77e-119 Length: 684  
 Score: 1159.00 Matches: 227  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 6 Gaps: 0

US-10-718-311-8 (1-227) x ABK47712 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
 DB 1 ATGGCTTCCTCTGTCAATTTCTTCAGCAGCTGTTCACACAGCGAGCAATGTTACACAAAGCT 60  
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
 DB 61 AGCATGGTTGCACCTTTCCTGGTCTCAATCTTCAGCCACCTTTCCTCTGTACAAAGAG 120  
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60  
 DB 121 CAAACACCTTGACATCACTTCCATTTGTAGCAATGGTGGAAAGATTAGCTGCATGCAGGTG 180  
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgGlyCysLysGluIle 80  
 DB 181 TGGCATATGTACACCCCGGTTAACGCAACTGCGTGGCTGCTGCTATTTGTAAGAGATC 240  
 QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100  
 DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATGACAAACGT 300  
 QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
 DB 301 TTTGAACAGCAGGAAACCGGTAAAGCGTGAAGATGATCCGCAAGGGTTTGTGAGCAG 360  
 QY 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTrpTrpLeuArgGlu 140  
 DB 361 AATGAATCCCCGAAAGTCCGCTGCTGCGCAAGAGTCTGTTACTGGTTACGGTAA 420  
 QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160  
 DB 421 ATTTTGTATGTCCGATGGTGNACCGTGGCTTGGCGTCTGACCGTCTGCTGTGTCA 480  
 QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTrp 180  
 DB 481 ACGTTAAGCGGCGCGAGCTGGCTGTACAAAATTTGGTAAACCGCGTTAGACGCTAT 540  
 QY 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGluLeu 200  
 DB 541 CTGTTACATCATCTGCATTAACCCGGGACTTTATTGAGATAGCGCGTATGCCGGGCTG 600  
 QY 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220  
 DB 601 TGGGGCGGACGTTCCCGCTTCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGACTGTTT 660

QY 221 LeuProAlaSerProLeuTyr 227  
 DB 661 TTACCGCGCTCACCGTTGTAC 681

#### RESULT 3

ADC25988

ID ADC25988 standard; DNA; 684 BP.

XX

AC ADC25988;

XX 18-DEC-2003 (first entry)

XX Tomato/Escherichia coli chimeric TP-CPL DNA.

XX UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;

KW liquid crystal polymer; LCP; methylparaben; preservative; food;

KW cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;

KW chorismate pyruvate lyase; plant.

XX

OS Chimeric.

OS Lycopersicon esculentum.

OS Escherichia coli.

XX

PH Key Location/Qualifiers

FT 1..684

FT /\*tag= a

FT /product= "Tomato/Escherichia coli chimeric TP-CPL

FT protein"

XX

XX WO2003066836-A2.

XX 14-AUG-2003.

XX 06-FEB-2003; 2003WO-US005863.

XX 07-FEB-2002; 2002US-0355511P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

XX WPI: 2003-767259/72.

XX P-PSDB; ADC25989.

XX New nucleic acid encoding UDP-glucosyltransferase, useful for preparing

PT cells that produce p-hydroxybenzoic acid glucose ester, also the new

PT enzymes.

XX Example 9; SEQ ID NO 41; 161pp; English.

XX The invention relates to a novel isolated nucleic acid that encodes a UDP

CC -glucosyltransferase. The method of the invention may be used to

CC transform microorganisms or green plant cells so that these produce a

CC higher level of high-volume chemicals or materials, such as pHBA (p-

CC hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a

CC monomer for liquid crystal polymers (LCPs) and starting material for

CC methylparaben. Methylparaben is a preservative commonly used in the food

CC and cosmetic industries. The encoded enzymes may be used for in vitro

CC production of these compounds and for identifying similar enzymes by

CC sequence comparison. The current sequence is that of the

CC tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA

XX of the invention.

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 5,77e-119 Length: 684  
 Score: 1159.00 Matches: 227  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 10 Gaps: 0

```
US-10-718-311-8 (1-227) x ADC25988 (1-684)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCTATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTCACTGGTCTCAATCTTCAGCCACTTTCCTGTTTACAAAGAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTCAACCCCGGTTAAACGCAACTGCGTGGCTGCTGCTATTTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGTGCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAGCT 300
QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTTGAACAGCAGGAGAAACGGTAAGCGGTGACGATGATCCGCGAAGGGTTTGTGCGAGCAG 360
QY 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAAATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGTTAGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
Db 421 ATTTTGTATGTGCCGATGGTGAACGGTGGCTTGGCGGTGCTGACGTTGTTCTGTGTC 480
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGCGAGTGGCGTTTACAAAATTTGGGTAAACCGCGTGGAGCGCTAT 540
QY 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTTCAATCATTCACATTAACCGGACTTTATTTGAGATAGGCGGTGATGCGGGCTG 600
QY 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTT 660
QY 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGGCTCACCGGTTGTAC 681
RESULT 4
ADV91622
ID ADV91622 standard; DNA; 684 BP.
XX
AC ADV91622;
XX
DT 10-MAR-2005 (first entry)
XX
DE Tomato rubisco small subunit precursor TP-E. coli CPL chimeric DNA.
XX
KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
KW HCHU; gene expression; hydroquinone; antiseptic; transgenic plant;
KW genetically engineered microorganism; antioxidant; antimicrobial agent;
KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
KW neoplasm; PHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid;
KW rubisco small subunit precursor; transit peptide;
KW chloroplast transit peptide; TP; chloroplast; gene; ds.
XX
OS Lycopersicon esculentum.
OS Escherichia coli.
OS Chimeric.
```

Db 181 TGGCATATGTCACACCCCGGTTAAAGCAACTCGCTGCGCTATTGTAAAGAGATC 240  
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100  
Db 241 CCTGCCCTGGATCGCAACTCGCTCGACTGGCTGTGCTGGAGGATTCATGACAAACGT 300  
Qy 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120  
Db 301 TTTGAACACAGCGAGGAAACCGTAAAGCGTACGATCGCGAAGGGTTTGTGAGCAG 360  
Qy 121 AsnGluIleProGluLeuLeuProLeuLeuProLysGluSerArgTrpTrpLeuArgGlu 140  
Db 361 ATGAATATCCCGAAGACTGCGCTGCTGCCGAAGAGTCTGTTACTGTTACGTGAA 420  
Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160  
Db 421 ATTTGTTATGTCGCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTCTCA 480  
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTrp 180  
Db 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATGGGTAAACGCCGTTAGACGCTAT 540  
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200  
Db 541 CTGTTACATCATCGACATTAAACCGGACTTTATTGAGATAGCGCTGATGCCGGCTG 600  
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220  
Db 601 TGGGGCGAGTTCCTCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTT 660  
Qy 221 LeuProAlaSerProLeuThr 227  
Db 661 TTACCGGCGTCAACGTTGTAC 681

## RESULT 5

ADV91636

ID ADV91636 standard; DNA; 3452 BP.

XX ADV91636;

XX 10-MAR-2005 (first entry)

XX C. parapsilosis pHB 1-H- E. coli CPL- A. thaliana UGT72B1 chimeric DNA.

KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;  
KW chorisinate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;  
KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;  
KW genetically engineered microorganism; antioxidant; antimicrobial agent;  
KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;  
KW neoplasm; pHBA 1-hydroxylase; pHBA 1-H; pHBA; para-hydroxybenzoic acid;  
KW p-hydroxybenzoic acid; UDP-glucosyltransferase; UGT72B1; gene; db.

XX Arabidopsis thaliana.

XX Candida parapsilosis.

XX Escherichia coli.

XX Chimeric.

XX Key Location/Qualifiers .

XX CDS 4. .501

XX /tag= a /product= "Escherichia coli CPL protein"

XX /tag= b /tag= .1980

XX /product= "Candida parapsilosis pHBA 1-H protein"

XX /transl\_except= (pos: 559. .561, aa:Trp)

XX CDS 2004. .3446

XX /tag= c

XX /product= "Arabidopsis thaliana UGT72B1 protein"

XX US2004261147-A1.

XX 23-DEC-2004.

XX

PF 16-JUN-2003; 2003US-00462162.  
XX  
PR 16-JUN-2003; 2003US-00462162.  
XX  
PA (MEYE/) MEYER K.  
PA (VIIT/) VIITANEN P V.  
PA (FLIN/) FLINT D.  
XX  
PI Meyer K, Viitanen PV, Flint D;  
XX  
XX WPI; 2005-057232/06.  
DR P-PSDB; ADV91605, ADV91619, ADV91631.  
XX  
PT Producing hydroquinone glucoside in a green plant comprises growing a  
PT green plant having nucleic acid fragments and chorisinate pyruvate lyase  
PT expression cassette.  
XX  
XX Example 7; SEQ ID NO 47; 70pp; English.  
XX  
XX The present invention relates to methods and materials to produce  
CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
CC in genetically modified green plants and microorganisms. The method  
CC relies upon transgenic plants or genetically modified microorganisms that  
CC produce increased levels of the initial substrate para-hydroxybenzoic  
CC acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce  
CC arbutin. Plants and microbes can be genetically engineered to produce  
CC high levels of pHBA either by functional expression of the bacterial  
CC protein chorisinate pyruvate lyase (CPL) or by the expression of 4-  
CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
CC present sequence is Candida parapsilosis pHBA 1-hydroxylase (pHB 1-H) -  
CC Escherichia coli chorisinate pyruvate lyase (CPL) and Arabidopsis thaliana  
CC UDP-glucosyltransferase (UGT72B1) chimeric DNA.  
XX  
SQ Sequence 3452 BP; 932 A; 730 C; 856 G; 934 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 6.56e-85 Length: 3452  
Score: 862.00 Matches: 166  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 74.4% Indels: 0  
DB: 14 Gaps: 0

US-10-718-311-8 (1-227) x ADV91636 (1-3452)

Qy 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIlePro 81  
Db 1 CATATGTACACCCCGCTTAACGCAACTGCGTGGCTGCTATTGTAAAGAGATCCCT 60  
Qy 82 AlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPhe 101  
Db 61 GCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGAGGATTCATGACAAACGTTTT 120  
Qy 102 GluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGlnAsn 121  
Db 121 GAACAGCAGGAGAAACCGGTAAAGCGGTACGATATCCGCAAGGGTGTTCGAGCAGAAT 180  
Qy 122 GluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIle 141  
Db 181 GAAATCCCCGAAGAACTGCCGCTGCGCCGAAGAGTCTGTTACTGGTTACGTGAAT 240  
Qy 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThr 161  
Db 241 TTGTTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTCTCTGTGTCACACG 300  
Qy 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181  
Db 301 TTAAGCGGCGGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGACGCTATCTG 360  
Qy 182 PheThrSerSerThrLeuThrArgaspPheIleGluIleGlyArgaspAlaGlyLeuTrp 201







```

QY 143 LeuCysAlaAspGlyGluProTrrpLeuAlaGlyArgThrValValProValSerThrLeu 162
DB 241 TTATGTGCGGATGGTGAACCGTGGCTTGGCGGTCTGACCGTCTTCTGTGTCAACGTTA 300
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
DB 301 AGCGGCGCGGAGCTGGCTTACAAAAATTGGGTAACAAACCGCTTAGGACGCTATCTGTTC 360
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
DB 361 ACATCATCGACATTAAACCGGACCTTATTGAGATAGCGCGGTGATGCCGGCTGTGGGG 420
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
DB 421 CGACGTTCCCGCTGGATTAGCGGTAACCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACG 480
QY 223 AlaSerProLeuTyr 227
DB 481 GCGTCACCGTTGTAC 495
RESULT 9
ABA91837
ID ABA91837 standard; DNA; 495 BP.
XX
AC ABA91837;
XX
DT 15-MAY-2002 (first entry)
XX
DE Escherichia coli chorismate pyruvate lyase gene coding region.
XX
KW Chorismate pyruvate lyase; CPL; enzyme; p-hydroxybenzoic acid;
KW transgenic plant; ubiC; gene; da.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..495
FT /tag= a
FT /partial
FT /product= "Chorismate pyruvate lyase"
FT /note= "the CDS does not include a stop codon"
XX
PN WO200194607-A2.
XX
XX 13-DEC-2001.
XX
XX 22-MAY-2001; 2001WO-US016661.
XX
XX 02-JUN-2000; 2000US-0209854P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Meyer K, Van Dyk DE, Viitanen PV;
XX
DR WPI; 2002-226795/28.
XX
XX P-PSDB; AAM50958.
XX
PT Producing para-hydroxy benzoic acid in green plant, comprises expression
PT of unique expression cassette containing gene encoding chorismate
PT pyruvate lyase operably linked to specific chloroplast targeting
PT sequence.
XX
PS Claim 4; Page 55; 60pp; English.
XX
CC The present sequence is that of the coding region of the Escherichia coli
CC strain W3110 ubiC gene in expression construct pET24a. The gene encodes
CC chorismate pyruvate lyase (CPL, see AAM50958). It was obtained by PCR
CC amplification of strain W3110 genomic DNA using primers (see ABA91839-40)
CC based on the published E. coli ubiC gene. The CPL open reading frame is
CC used in an expression cassette designed for the high-level production of
CC p-hydroxybenzoic acid (pHBA) in green plants. The expression cassette
CC comprises the CPL coding sequence operably linked to a promoter capable
CC of driving protein expression in higher plants. The cassette also has a

```

```

CC sequence encoding a chloroplast transit peptide, its natural cleavage
CC site, and a small portion of a transit peptide donor protein fused to the
CC N-terminus of CPL. The chloroplast targeting sequence targets the foreign
CC protein to the chloroplast and aids in its uptake into the organelle. The
CC cleavage site is unique to the transit peptide, and cleavage of the
CC chimeric protein encoded by the cassette at this site releases a novel
CC polypeptide that has full enzyme activity, comprising the mature CPL
CC enzyme and a small portion of transit peptide donor (see AAM50961). A
CC plant comprising the CPL expression cassette is claimed, and may be
CC soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
CC barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,
CC millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
CC component of liquid crystal polymers which have application in the
CC automotive, electrical and other industries
XX
SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 3.04e-85 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 6 Gaps: 0

US-10-718-311-8 (1-227) x ABA91837 (1-495)
QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLeuGluLeuProAla 82
DB 1 ATGTGCACACCCCGCTTAAACGCAACTGCGTGGCTATTGTAAAGAGATCCCTGCC 60
QY 83 LeuAspProGlnLeuLeuAspTrrpLeuLeuLeuGluAspSerMetThrLysArgPheGlu 102
DB 61 CTGGATCCGCACTGCTCGACTGGCTGTGCTGGAGGATTCATGACAAACGTTTGAA 120
QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
DB 121 CAGCAGGAAAAACCGTAAGCGTGCATGATCGCGAAAGGTTTGTGAGCAGAAATGAA 180
QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrrpLeuArgGluLeu 142
DB 181 ATCCCCGAAGAACTGCGCGTCTGCCGAAGAGTCTCGTTACTGTTACGTGAAATTTTG 240
QY 143 LeuCysAlaAspGlyGluProTrrpLeuAlaGlyArgThrValValProValSerThrLeu 162
DB 241 TTATGTCCGATGTTGAACCGTGGCTTGGCGTGTACCGTCTTCTGTCTCAACGTTA 300
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
DB 301 AGCGGCGCGGAGCTGGCGTTACAAAAATTGGGTAACAAACCGCTTAGGACGCTATCTGTTC 360
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
DB 361 ACATCATCGACATTAAACCGGACCTTATTGAGATAGCGCGGTGATGCCGGCTGTGGGG 420
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
DB 421 CGACGTTCCCGCTGGATTAAACCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACG 480
QY 223 AlaSerProLeuTyr 227
DB 481 GCGTCACCGTTGTAC 495
RESULT 10
ABA47709
ID ABA47709 standard; DNA; 495 BP.
XX
AC ABA47709;
XX
DT 18-JUN-2002 (first entry)
XX
DE DNA sequence of ORF for CPL in pET24a E. coli expression construct.
XX

```



KW SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester;  
 KW sinapoylglucose:malate sinapoyltransferase; glycosylated aromatic acid;  
 KW malate conjugated aromatic acid; polymer synthesis; ubiC; CPL; gene;  
 KW carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;  
 KW pET24a; ds.  
 XX Escherichia coli.  
 FH Key Location/Qualifiers  
 FT CDS 1..495  
 FT /\*tag= a  
 FT /partial  
 FT /product= "CPL"  
 FT /note= "This sequence lacks a stop codon"  
 XX WO200204653-A2.  
 PN 17-JAN-2002.  
 XX 05-JUL-2001; 2001WO-US021283.  
 XX 07-JUL-2000; 2000US-0216615P.  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PI Flint D, Meyer K, Viitanen PV;  
 XX WPI; 2002-303779/34.  
 XX P-PSDB; AAU77941.  
 DR Producing aromatic acid conjugates, involves contacting glycosylated  
 PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the  
 PT presence of sinapoylglucose:malate sinapoyltransferase.  
 XX Example; Page 68; 72pp; English.  
 XX The present invention relates to the isolation of Arabidopsis thaliana  
 CC gene (SNG1, sinapoylglucose accumulator 1) encoding  
 CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the  
 CC substitution of a glucose moiety on a glycosylated aromatic acid with a  
 CC malate moiety to form a malate conjugated aromatic acid. The enzyme is  
 CC useful for producing malate conjugated aromatic acids, carboxylic acid  
 CC conjugated aromatic acids or aromatic esters. The malate conjugated  
 CC aromatic acids are useful in the synthesis of various polymers. The  
 CC present sequence representing the ORF (open reading frame) for chorismate  
 CC pyruvate lyase (CPL) in pET24a Escherichia coli expression construct is  
 CC described in the examples of the present invention  
 XX SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,04e-85 Length: 495  
 Score: 854.00 Matches: 165  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 73.7% Indels: 0  
 DB: 6 Gaps: 0

US-10-718-311-8 (1-227) x ABK47709 (1-495)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysIysGluileProAla 82  
 DB 1 ATGTACACCCCGCGTAAACGCAACTGCGTGGCGCTATTGTAAAGAGATCCCTGCC 60  
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrIysArgPheGlu 102  
 DB 61 CTGGATCCCAACTGCTGACCTGGCTGTGGAGGATCCATGACAAAGATTGAA 120  
 QY 103 GlnGlnGlyLeuThrValSerValThrMetIleArgGluGlyPheValGlnGlnAenGlu 122  
 DB 121 CAGCAGGGAAAAACGGTAAAGCGTACGATGATCCGCGAAGGGTTGTCGAGCAGAAATGAA 180  
 QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluileLeu 142

Db 181 ATCCCGAAGAACTCCCGCTGCTCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTG 240  
 QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162  
 Db 241 TTATGTCCGATGGTGAACCGTGGCTTGCCGTCGTACCGTCTCTCTGTCACAGTTA 300  
 QY 163 SerGlyProGluLeuAlaLeuGlnIysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
 Db 301 AGCGGCGCGAGCTGGCGTTACAAAAATTGGGTAAACGCCGCTTAGGACGCTATCTGTTTC 360  
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
 Db 361 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGGCCGCGTATGCCGGCTGTGGGG 420  
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuLeuLeuLeuLeuPro 222  
 Db 421 CGACGTTCCCGCTCGGATTAAAGCGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480  
 QY 223 AlaSerProLeuTyr 227  
 Db 481 GCGTCACCGTTGTAC 495

RESULT 11  
 ADC25984  
 ID ADC25984 standard; DNA; 495 BP.  
 XX  
 AC ADC25984;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Escherichia coli chorismate pyruvate lyase DNA.  
 KW UDP-glucosyltransferase; PHBA; p-hydroxybenzoic acid ester glucoside;  
 KW liquid crystal polymer; LCP; methylparaben; preservative; food;  
 KW cosmetic industry; ds; gene; ubiC; CPL; chorismate pyruvate lyase.  
 XX Escherichia coli.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 1..495  
 FT /\*tag= a  
 FT /partial  
 FT /product= "Escherichia coli chorismate pyruvate lyase  
 FT protein"  
 FT /note= "No stop codon"  
 XX WO2003066836-A2.  
 XX  
 PD 14-AUG-2003.  
 XX  
 PF 06-FEB-2003; 2003WO-US005863.  
 XX  
 PR 07-FEB-2002; 2002US-0355511P.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Meyer K, Van Dyk DE, Viitanen PV;  
 XX  
 DR WPI; 2003-767259/72.  
 XX P-PSDB; ADC25985.  
 PT New nucleic acid encoding UDP-glucosyltransferase, useful for preparing  
 PT cells that produce p-hydroxybenzoic acid glucose ester, also the new  
 PT enzymes.  
 XX  
 PS Example 9; SEQ ID NO 37; 161pp; English.  
 XX

The invention relates to a novel isolated nucleic acid that encodes a UDP-glucosyltransferase. The method of the invention may be used to transform microorganisms or green plant cells so that these produce a higher level of high-volume chemicals or materials, such as PHBA (p-hydroxybenzoic acid) ester glucoside which is an intermediate for PHBA, a

CC monomer for liquid crystal polymers (LCPs) and starting material for  
 CC methylparaben. Methylparaben is a preservative commonly used in the food  
 CC and cosmetic industries. The encoded enzymes may be used for in vitro  
 CC production of these compounds and for identifying similar enzymes by  
 CC sequence comparison. The current sequence is that of the *Escherichia coli*  
 CC chorismate pyruvate lyase (CPL; UbiC) DNA of the invention.

XX SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 3.04e-85 Length: 495  
 Score: 854.00 Matches: 165  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 73.7% Indels: 0  
 DB: 10 Gaps: 0

US-10-718-311-8 (1-227) x ADC25984 (1-495)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
 DB 1 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 60  
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
 DB 61 CTGGATCCGCAACTGCTGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 120  
 QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
 DB 121 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 180  
 QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
 DB 181 ATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAATTTTG 240  
 QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162  
 DB 241 TTATGTGCGGATGGTGAACCGTGGCTTCCGCGTGTACCGTCTGTTCAACGTTA 300  
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182  
 DB 301 AGCGGCGGAGCTGCGCTTACAAATTTGGTAAACCGCGTTAGGACGCTATCTGTTTC 360  
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
 DB 361 ACATCATCGACATTAAACCGGACTTTTATGAGATAGCCCGTGTATGCCGGCTGTGGGG 420  
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
 DB 421 CGACGTTCCCGCTCGGATTAAAGCGGTAACCGCTGTTGCTTAACAGAACTGTTTTACCG 480  
 QY 223 AlaSerProLeuTyr 227  
 DB 481 CGGTACCGTGTATC 495

#### RESULT 12

ADV91618

ID ADV91618 standard; DNA; 498 BP.

XX AC ADV91618;

XX DT 10-MAR-2005 (first entry)

XX DE *Escherichia coli* chorismate pyruvate lyase (CPL) DNA.

XX Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;  
 KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;  
 KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;  
 KW genetically engineered microorganism; antioxidant; antimicrobial agent;  
 KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;  
 KW neoplasm; pHEA; para-hydroxybenzoic acid; p-hydroxybenzoic acid; gene;  
 KW ds.

OS *Escherichia coli*.  
 XX Key Location/Qualifiers  
 FH 1. .498  
 FT /\*tag= a  
 FT /product= "Escherichia coli CPL protein"  
 XX US2004261147-A1.  
 PN 23-DEC-2004.  
 XX 16-JUN-2003; 2003US-00462162.  
 XX 16-JUN-2003; 2003US-00462162.  
 PR (MEYE/) MEYER K.  
 PA (VIIT/) VIITANEN P V.  
 PA (FLIN/) FLINT D.  
 XX Meyer K, Viitanen PV, Flint D;  
 PI WPI; 2005-057232/06.  
 DR P-ESDB; ADV91619.

XX Producing hydroquinone glucoside in a green plant comprises growing a  
 green plant having nucleic acid fragments and chorismate pyruvate lyase  
 PT expression cassette.  
 PT  
 XX Example 4; SEQ ID NO 29; 70pp; English.  
 XX The present invention relates to methods and materials to produce  
 CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
 CC in genetically modified green plants and microorganisms. The method  
 CC relies upon transgenic plants or genetically modified microorganisms that  
 CC produce increased levels of the initial substrate para-hydroxybenzoic  
 CC acid (p-hydroxybenzoic acid; pHEA) in a biosynthetic pathway to produce  
 CC arbutin. Plants and microbes can be genetically engineered to produce  
 CC high levels of pHEA either by functional expression of the bacterial  
 CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-  
 CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
 CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
 CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
 CC present sequence is *Escherichia coli* chorismate pyruvate lyase (CPL) DNA.  
 CC This sequence is found in the *E. coli* expression vector construct  
 CC pET24a(+).

XX SQ Sequence 498 BP; 114 A; 122 C; 141 G; 121 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.06e-85 Length: 498  
 Score: 854.00 Matches: 165  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 73.7% Indels: 0  
 DB: 14 Gaps: 0

US-10-718-311-8 (1-227) x ADV91618 (1-498)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82  
 DB 1 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 60  
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102  
 DB 61 CTGGATCCGCAACTGCTGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 120  
 QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122  
 DB 121 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 180  
 QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142  
 DB 181 ATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAATTTTG 240

QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162  
 DB 241 TTATGTGCCGATGTCGAACCGTGGCTTGCCTGCGTACCGTCTGCTGTCAACGTTA 300  
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgThrLeuPhe 182  
 DB 301 AGCGGGCCGAGCTGGCGTTACAAAATGGGTAAACCGCGTTAGACGCTATCTGTTT 360  
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202  
 DB 361 ACATCATCGACATTAACCGCGGACTTATTGAGATAGCGCGTGTATGCCGGCTGTGGGG 420  
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222  
 DB 421 CGACGTTCCCGCTGGGATTAAGCGGTAAACCGCTGTGTGTAACAGAACTGTTTTTACCG 480  
 QY 223 AlaSerProLeuTyr 227  
 DB 481 CGGTACCGTTGTAC 495  
 RESULT 13  
 AAT29820  
 ID AAT29820 standard; DNA; 2000 BP.  
 AC AAT29820;  
 XX  
 DT 17-JAN-1997 (first entry)  
 XX  
 DE E. coli ubiquinone biosynthetic enzyme genes ubiC and ubiA.  
 XX  
 KW Ubiquinone-10; biosynthetic enzyme; ubiC; ubiA; gene;  
 KW photosynthetic bacteria; recombinant production; plasmid pRSFAC;  
 KW glutamate synthase; gene promoter; Rhodobacter capsulatus; transformant;  
 KW MC9R/pRSFAC; ss.  
 XX  
 OS Escherichia coli.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 380..877  
 FT /\*tag= a  
 FT /label= ubiC  
 FT CDS 890..1782  
 FT /\*tag= b  
 FT /label= ubiA  
 FT repeat\_unit 1868..1887  
 FT /\*tag= c  
 FT /rpt\_type= inverted  
 FT repeat\_unit 1892..1911  
 FT /\*tag= d  
 FT /rpt\_type= inverted  
 FT FT  
 PN JP08107789-A.  
 XX  
 PD 30-APR-1996.  
 XX  
 PF 13-OCT-1994; 94JP-00273071.  
 XX  
 PR 13-OCT-1994; 94JP-00273071.  
 XX  
 PA (ALPHA-) ALPHA SHOKUIN KK.  
 XX  
 WIPI; 1996-262599/27.  
 DR P-PSDB; AAR97746, AAR97747.  
 XX  
 PT Genes coding for ubiquinone biosynthetic enzymes - useful for prodn. of  
 PT ubiquinone-10 by transformed photosynthetic bacteria.  
 XX  
 PS Claim 1; Fig 2; 11pp; Japanese.  
 XX  
 CC The present sequence comprises the ubiquinone biosynthetic enzyme genes  
 CC ubiC and ubiA, which were cloned from the chromosomal DNA of E. coli  
 CC (Kohara map phage DNA bank IF8 (634); Cell 50, 495-508 (1987)). A large

CC amt. of ubiquinone-10 can be produced by culturing photosynthetic  
 CC bacteria transformed with the novel plasmid pRSFAC, which contains the  
 CC ubiC and ubiA genes and the glutamate synthase gene promoter,  
 CC specifically the Rhodobacter capsulatus transformant MC9R/pRSFAC  
 XX  
 SQ Sequence 2000 BP; 444 A; 462 C; 538 G; 556 T; 0 U; 0 Other;  
 Alignment Scores: 2.98e-84 Length: 2000  
 Pred. No.: 853.00 Matches: 165  
 Score: 97.7% Conservative: 2  
 Percent Similarity: 97.7% Mismatches: 4  
 Best Local Similarity: 96.5% Indels: 0  
 Query Match: 73.6% Gaps: 0  
 DB: 2  
 US-10-718-311-8 (1-227) x AAT29820 (1-2000)  
 QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76  
 DB 362 TGTAAACGGAGAGTTCGGCATGTACACCCCGGCTTAACGCACTGCGTGGCTGCGCTAT 421  
 QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96  
 DB 422 TGTAAAGAGATCCCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTCGTGGAGGATTCC 481  
 QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
 DB 482 ATGACAAAACGTTTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGCAGATGATGCGCGAAGG 541  
 QY 117 PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 DB 542 TTGTGCGAGCAGAAATGAAATCCCGAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTAC 601  
 QY 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156  
 DB 602 TGGTACGTGAAATTTTGTATGTCCGATGGTGAACCGTGGCTTCCCGTCCGCGTCCGCGTC 661  
 QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176  
 DB 662 GTTCTCTGTCAACGTTAAGCGGCGGAGCTGGCGCTTACAAAAATTTGGGTAAACGCGC 721  
 QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196  
 DB 722 TTAGGACGCTATCTGTTCATCATCATGACATTAAACCCGGGACTTTATTGAGATAGGCCGT 781  
 QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216  
 DB 782 GATCGCGGCTGTGGGGCGGAGCTTCCCGCTCGGATTAAAGCGGTAACCGCTGTTGCTA 841  
 QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227  
 DB 842 ACAGAACTGTTTTTACCGCGCTCACCGTTGTAC 874  
 RESULT 14  
 ADV91643  
 ID ADV91643 standard; DNA; 1971 BP.  
 XX  
 AC ADV91643;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX  
 DE C. parapsilosis pH8 1-H- E. coli CPL chimeric DNA.  
 XX  
 KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;  
 KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;  
 KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;  
 KW genetically engineered microorganism; antioxidant; antimicrobial agent;  
 KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cycostatic;  
 KW neoplasm; pHBA 1-hydroxylase; PHBA; para-hydroxybenzoic acid;  
 KW p-hydroxybenzoic acid; gene; ds.  
 XX  
 OS Candida parapsilosis.  
 OS Escherichia coli.

OS Chimeric.  
 XX Key Location/Qualifiers  
 FT CDS 4..501 /tag= a  
 FT /product= "Escherichia coli CPL protein"  
 FT /transl\_except= (pos: 46..48, aa:Cys)  
 FT 522..1561  
 FT /tag= b  
 FT /product= "Candida parapsilosis pHEA 1-H protein"  
 FT /transl\_except= (pos: 540..542, aa:Trp)  
 XX US2004261147-A1.  
 XX 23-DEC-2004.  
 XX 16-JUN-2003; 2003US-00462162.  
 XX 16-JUN-2003; 2003US-00462162.  
 XX (MEYE/) MEYER K.  
 PA (VIIT/) VIITANEN P V.  
 PA (FLIN/) FLINT D.  
 XX Meyer K, Viitanen PV, Flint D;  
 PI WPI; 2005-057232/06.  
 DR P-PSDB; ADV91605, ADV91619.  
 XX Producing hydroquinone glucoside in a green plant comprises growing a  
 PT green plant having nucleic acid fragments and chorismate pyruvate lyase  
 PT expression cassette.  
 XX Example 7; SEQ ID NO 54; 70pp; English.  
 XX The present invention relates to methods and materials to produce  
 CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
 CC in genetically modified green plants and microorganisms. The method  
 CC relies upon transgenic plants or genetically modified microorganisms that  
 CC produce increased levels of the initial substrate para-hydroxybenzoic  
 CC acid (p-hydroxybenzoic acid; pHEA) in a biosynthetic pathway to produce  
 CC arbutin. Plants and microbes can be genetically engineered to produce  
 CC high levels of pHEA either by functional expression of the bacterial  
 CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-  
 CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
 CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
 CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
 CC present sequence is Candida parapsilosis pHEA 1-hydroxylase (pHB 1-H) and  
 CC Escherichia coli chorismate pyruvate lyase (CPL) chimeric DNA.  
 XX Sequence 1971 BP; 550 A; 383 C; 484 G; 554 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 4,87e-84 Length: 1971  
 Score: 851.00 Matches: 165  
 Percent Similarity: 99.4% Conservative: 0  
 Best Local Similarity: 99.4% Mismatches: 1  
 Query Match: 73.4% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-718-311-8 (1-227) x ADV91643 (1-1971)  
 QY 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIlePro 81  
 Db 1 CATATGTACACCCCGGTTAAGCAACTCGTGGCTATTTAAAGAGATCCCT 60  
 QY 82 AlaLeuAppProGlnLeuLeuAspTrpLeuLeuGluLeuAspSerMetThrLysArgPhe 101  
 Db 61 GCGCTGGATCCGCAACTGCTCGAGTGGCTTTGGCTGGAGGATTCACACAAACGTTTT 120  
 QY 102 GluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsn 121  
 Db 121 GAACAGCAGGGGAAAACGGTAAGCGTGACGATGATCCCGAAGGGTTTGTCTGAGCAGAT 180

QY 122 GluIleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIle 141  
 Db 181 GAAATCCCGAAGAACTGCGCTGTCGCGAAAGAGTCTCGTTACTGGTTAGTGAATT 240  
 QY 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThr 161  
 Db 241 TTGTTATGTGCGCATGGTGAACCGTGGCTTGCCTGCTACCGTCTTCTCTGTCAACG 300  
 QY 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181  
 Db 301 TTAAGCGGCGGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGGAGCTATCTG 360  
 QY 182 PheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrp 201  
 Db 361 TTCACATCATCGACATTAAACCGGACCTTTATGAGATAGCCGCTGATGCCGGCTGTGG 420  
 QY 202 GlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221  
 Db 421 GGGCGACGTTCCCGCTCGGATTAAAGCGTAAACCGCTGTTGCTAACAGAACTGTTTTTA 480  
 QY 222 ProAlaSerProLeuTyr 227  
 Db 481 CCGCGCTCACCGTTGTAC 498  
 RESULT 15  
 AAS90335  
 ID AAS90335 standard; cDNA; 1207 BP.  
 XX AAS90335;  
 AC AAS90335;  
 DT 13-FEB-2002 (first entry)  
 XX DNA encoding novel human diagnostic protein #26139.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.  
 FN WO200175067-A2.  
 XX 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR P-PSDB; ABG26148.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 26139; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 1207 BP; 316 A; 265 C; 288 G; 338 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	4.03e-80	Length:	1207
Score:	813.00	Matches:	164
Percent Similarity:	96.0%	Conservative:	2
Best Local Similarity:	94.8%	Mismatches:	5
Query Match:	70.1%	Indels:	2
DB:	5	Gaps:	0

US-10-718-311-8 (1-227) x AAS90335 (1-1207)

QY	57	CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr	76
DB	634	TGTAACGGAGAGTTCGGCATGTACACCCCGGTTAAACGCAACTGCGCTGCGCTAT	693
QY	77	CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer	96
DB	694	TGTAAGAGATCCCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTGTGGAGGATCC	753
QY	97	MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly	116
DB	754	ATGACAAAACGTTTGAACAGCAGGGAACAAACGGTAAGCGTGACGATGATCCGCGAAGG	813
QY	117	PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr	136
DB	814	TTTGTGAGCAGAGTAATAATCCCGAAGAACTGCCGCTGCTCCGAAAGAGTCTCGTTAC	873
QY	137	TrpLeuArgGluIleLeu-LeuCysAlaAspGlyGluProTrpLeuAla-GlyArgThrV	156
DB	874	TGGTTACGTGAATTTTGTGTATGTGCCATGGCGAACCGCGGCTTGCCGGGTCTGACCG	933
QY	156	alValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrP	176
DB	934	TCGTTCTCTGTCAACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATGGGTAAACGC	993
QY	176	roLeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyA	196
DB	994	CGTTAGGACGCTATCTGTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCC	1053
QY	196	rgAspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuL	216
DB	1054	GTGATCCCGGCTGTGGGGGCGACGTTCCCGCTTCGATTAGCGGTAAACCGCTGTTC	1113
QY	216	euThrGluLeuPheLeuProAlaSerProLeuTyr	227
DB	1114	TACAGAACTGTTTTTACCGGGGCTCACCGTGTGTAC	1148

Search completed: July 11, 2006, 21:50:40  
Job time : 637.516 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

Run on: July 11, 2006, 19:46:18 ; Search time 4967.29 Seconds  
(without alignments)  
4383.493 Million cell updates/sec

Title: US-10-718-311-8  
Perfect score: 1159  
Sequence: 1 MASSVISAATRSNVTOA.....RLSGKPLLTLELPASPLY 227

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 6366136 seqs, 31973710525 residues  
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSWEB spool/US10718311/runat.11072006.111520.5171/app.query.fasta.1  
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03p  
-USER=US10718311 @CGN 1.1 8663 @runat.11072006.111520.5171 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.env.\*  
2: gb.pat.\*  
3: gb.ph.\*  
4: gb.pl.\*  
5: gb.pr.\*  
6: gb.ro.\*  
7: gb.sts.\*  
8: gb.sy.\*  
9: gb.un.\*  
10: gb.vi.\*  
11: gb.ov.\*  
12: gb.htg.\*  
13: gb.in.\*  
14: gb.om.\*  
15: gb.ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1159	100.0	684	2	AR428736 Sequence
2	1159	100.0	684	2	AR455412 Sequence
3	1159	100.0	684	2	AX329368 Sequence

4	1159	100.0	684	2	AX382258 Sequence
5	856	73.9	854	2	AR144788 Sequence
6	856	73.9	1498	15	M92668 Escherichia
7	856	73.9	2000	2	E11274 ubiC and ub
8	856	73.9	2000	15	M93136 Escherichia
9	856	73.9	2000	15	M93413 Escherichia
10	856	73.9	2034	15	X57434 E.coli ubiC
11	856	73.9	2348	15	X66619 E.coli gene
12	856	73.9	6641	2	AR144789 Sequence
13	856	73.9	110000	15	Continuation (43 o
14	856	73.9	110000	15	Continuation (43 o
15	856	73.9	176195	15	U00006 E. coli chr
16	854	73.7	495	2	A47932 Sequence 1
17	854	73.7	495	2	AR428733 Sequence
18	854	73.7	495	2	AR455409 Sequence
19	854	73.7	495	2	AX329364 Sequence
20	854	73.7	495	2	AX382254 Sequence
21	852	73.5	110000	15	Continuation (52 o
22	848	73.2	110000	15	Continuation (51 o
23	848	73.2	110000	15	Continuation (52 o
24	845	72.9	110000	15	Continuation (44 o
25	845	72.9	110000	15	Continuation (35 o
26	845	72.9	110000	15	Continuation (42 o
27	843	72.7	1396	15	DO087228 Escherich
28	841	72.6	110000	15	Continuation (43 o
29	837	72.2	110000	15	Continuation (45 o
30	835	72.0	110000	15	Continuation (48 o
31	800	69.0	1918	2	E07849 gDNA encodi
32	725	62.6	23880	15	AE008898 Salmonell
33	725	62.6	110000	15	Continuation (42 o
34	725	62.6	110000	15	Continuation (43 o
35	721	62.2	110000	15	Continuation (43 o
36	721	62.2	110000	15	Continuation (44 o
37	721	62.2	234050	15	AL627282 Salmonell
38	673	58.1	645	2	AR384505 Sequence
39	438.5	37.8	110000	15	Continuation (5 of
40	437.5	37.7	10668	15	AE013658 Versinia
41	437.5	37.7	214050	15	AJ414142 Versinia
42	437.5	37.7	290803	15	AE017128 Versinia
43	424.5	36.6	110000	15	Continuation (7 of
44	417	36.0	349107	15	BX571873 Phototrab
45	408.5	35.2	110000	15	Continuation (13 o

ALIGNMENTS

RESULT 1	AR428736	Sequence 16 from patent US 6642036.	684 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR428736	Sequence 16 from patent US 6642036.	684 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	AR428736	Sequence 16 from patent US 6642036.	684 bp	DNA	linear	PAT 18-DEC-2003
ACCESSION	AR428736	Sequence 16 from patent US 6642036.	684 bp	DNA	linear	PAT 18-DEC-2003
VERSION	AR428736.1	GI:40188466	684 bp	DNA	linear	PAT 18-DEC-2003
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 684)					
AUTHORS	Flint, D., Meyer, K. and Viitanen, P.					
TITLE	Sinapoylglucose:malate sinapoyltransferase form malate conjugates from benzoic acid glucosides					
JOURNAL	Patent: US 6642036-A 16 04-NOV-2003;					
	E. I. Du Pont de Nemours and Company; Wilmington;					
	WOX;					
FEATURES	Location/Qualifiers					
source	1. .684					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						
Alignment Scores:						
Pred. No.:	1.12e-110					684
Score:	1159.00					227
Percent Similarity:	100.0%					0

Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

  

US-10-718-311-8 (1-227) x AR428736 (1-684)			
Qy	1	MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla	20
Db	1	ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGCCACACGAGCAATGTTACACAAGCT	60
Qy	21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40
Db	61	AGCATGGTGGACCTTCTCAAAATCTTCAGCAGCTGTGCCACACGAGCAATGTTACACAAGCT	60
Qy	41	GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60
Db	121	CAAAACCTTGACATCCTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG	180
Qy	61	TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle	80
Db	181	TGGCATATGTACACCCCGGTTAAACGCACTGCGTGCCTATTGTAAGAGATC	240
Qy	81	ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg	100
Db	241	CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTGCTGGAGGATTCATGACAAAACGT	300
Qy	101	PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln	120
Db	301	TTTGAACAGCAGGGAACCGTAAGCGTGACGATGATCCCGAAGGGTTGTGCGAGCAG	360
Qy	121	AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu	140
Db	361	AATGAATCCCGAAGAACTGCGCTGTGCGCAAAAGAGTCTCGTTACTGGTTACGTGTA	420
Qy	141	IleLeuLeuCysAlaAspGlyProTrpLeuAlaGlyArgThrValProValSer	160
Db	421	ATTTTGTATGTGCGCATGGTAACCGTGGCTGCCGTAAGAGTCTCGTTACTGGTTACGTGTA	480
Qy	161	LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu	200
Db	541	CTGTTTACATCATCGACATTAACCCGGGACITTTATTGAGATAGCCGCGTATGCCGGCTG	600
Qy	201	TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPhe	220
Db	601	TGGGGCGCAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT	660
Qy	221	LeuProAlaSerProLeuTyr	227
Db	661	TTACCGGCGTCACCGTTGTAC	681

  

RESULT 2	AR455412	684 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR455412	Sequence 7 from patent US 5683231.			
DEFINITION	AR455412				
ACCESSION	AR455412.1	GI:42689960			
VERSION	AR455412.1	GI:42689960			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 684)				
AUTHORS	Meyer, K., Viitanen, P.V. and Van Dyk, D.E.				
TITLE	High level production of p-hydroxybenzoic acid in green plants				
JOURNAL	Patent: US 5683231-A 7 27-JAN-2004;				
	E. I. du Pont de Nemours and Company; Wilmington, DE				
FEATURES	Location/Qualifiers				
source	1..684				
	/organism="unknown"				
	/mol_type="genomic DNA"				

  

RESULT 3	AX329368	684 bp	DNA	linear	PAT 08-JAN-2002
LOCUS	AX329368	Sequence 7 from Patent WO0194607.			
DEFINITION	AX329368				
ACCESSION	AX329368				
VERSION	AX329368.1	GI:18102367			
KEYWORDS	.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	1				
AUTHORS	Meyer, K., van Dyk, D.E. and Viitanen, P.V.				
TITLE	High level production of p-hydroxybenzoic acid in green plants				



```
JOURNAL Patent: WO 0194607-A 7 13-DEC-2001;
FEATURES E.I. Dupont De Nemours (US)
SOURCE Location/Qualifiers
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic CPL"

ORIGIN
Alignment Scores:
Pred. No.: 112e-110 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x AX329368 (1-684)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTTCATCGTCTCAAAATCTTCAGCCACCTTCCTCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAACCTTGACATCACTTCATTTGCTAGCAATGTGGGAAGTAGTGCATGCAAGGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB 181 TGGCATATGTCACACCCCGGTTAACGCAACTCGGTGCGCTGCTATTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTTCGGAGGATTCATGACAAACGT 300
QY 101 PheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
DB 301 TTTGAACACAGCGGAAAAACGGTAGCGTAGCATGATCCGCGAGGGTTGTTCGAGCAG 360
QY 121 AsnGluIleProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
DB 361 AATGAAATCCCCGAACAACCTGCGCTGCTCCGCAAGAGTCTCGTTACTGTTACGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160
DB 421 ATTTTGTATGTCCGATGGTGAACCGTGGCTTGCCTGCTGCTACCGCTCTCTGTGTC 480
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
DB 481 ACGTTAAGCGGCGCGAGCTGGCTGTTACAAAAATTTGGGTAAAAACCGCGTTAGACGCTAT 540
QY 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
DB 541 CTGTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTATGCCGGGCTG 600
QY 201 TrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
DB 601 TGGGGCGAGCTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
QY 221 LeuProAlaSerProLeuTyr 227
DB 661 TTACCGCGCTCACCGTTGTAC 681

RESULT 4
AX382258 684 bp DNA linear PAT 18-MAR-2002
LOCUS AX382258
DEFINITION Sequence 16 from Patent WO0204653.
ACCESSION AX382258
```



```

Pred. No.: 9,36e-79 Length: 1498
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservatives: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0

US-10-718-311-8 (1-227) x ECOUBIA (1-1498)

QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
   |||
Db 729 TGTAAACGAGAGTTCGGCATGTACACCCCGCGTTAAACGCAACTGCGTGGCGTAT 788

QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
   |||
Db 789 TGTAAAGAGATCCCTCCCTCGGATCCGCAACTCTCGACTGGCTGTGTGCTGGAGGATTC 848

QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
   |||
Db 849 ATGACAAAACGTTTGNACAGCAGGGGNAACCGTAAAGCGTGACGATGATCCCGGAGGG 908

QY 117 PheValGlnGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
   |||
Db 909 TTTGTGCGAGAGATGAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 968

QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
   |||
Db 969 TGGTTACGTGAAATTTGTTATGTCGGATGTTGAACCGTGGTTCGCGTCTGATCCGTC 1028

QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
   |||
Db 1029 GTTCTGTGTCAAGTTAAGCGGCGGAGCTGCGGTTTACAAAATTTGGTAAACGCCG 1088

QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
   |||
Db 1089 TTAGGACGCTATCTGTTACATCATCGACATTAAACCGGGACITTTATTGAGATAGCCGT 1148

QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
   |||
Db 1149 GATCGCGGGCTGTGGGGCGAGCTTCGCGCTCGGATTAAAGCGGTAACCGTGTGTGCTA 1208

QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
   |||
Db 1209 ACAGAACTGTTTACCGGCGTCACCGTGTGTAC 1241

RESULT 7
E11274
LOCUS E11274 2000 bp DNA linear PAT 29-SEP-1997
DEFINITION ubiC and ubiA gene.
ACCESSION E11274
VERSION E11274.1 GI:22024916
KEYWORDS JP 1996107789-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE Matsuda,H., Kawamuki,M. and Nakagawa,T.
AUTHORS PRODUCTION OF UBIQUINONE-10.
JOURNAL Patent: JP 1996107789-A 1 30-APR-1996;
ALPHA- SHOKUHIN KK
OS Escherichia coli
PN JP 1996107789-A/1
PD 30-APR-1996
PF 13-OCT-1994 JP 1994273071
PI MATSUDA HIDEYUKI, KAWAMUKI MAKOTO, NAKAGAWA TSUYOSHI PC
(C12N1/21, C12N15/09,A61K31/12,A61K31/12,C07H21/04,C12N1/21,C12P7/66, PC
PC C12R1/01), (C12P7/66,C12R1/01);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH

```

```

FT source 1..2000
FT /organism='Escherichia coli'
FT CDS 380..877
FT /product='UbiC protein'
FT CDS 890..1762
FT /product='UbiA protein'.
FEATURES
    Location/Qualifiers
    source 1..2000
    /organism="Escherichia coli"
    /mol_type="genomic DNA"
    /db_xref="taxon:562"
ORIGIN
Alignment Scores:
Pred. No.: 1,33e-78 Length: 2000
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservatives: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 2 Gaps: 0
US-10-718-311-8 (1-227) x E11274 (1-2000)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
   |||
Db 362 TGTAAACGAGAGTTCGGCATGTACACCCCGCGTTAAACGCAACTGCGTGGCGTAT 421

QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
   |||
Db 422 TGTAAAGAGATCCCTCCCTCGGATCCGCAACTGCTCGACTGGCTGTGTGCTGGAGGATTC 481

QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
   |||
Db 482 ATGACAAAACGTTTGAACAGCAGGGGAAAAACCGTAAAGCGTGCAGATGATCCCGAAGGG 541

QY 117 PheValGlnGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
   |||
Db 542 TTTGTGACAGAGATGAAATCCCGAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTAC 601

QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
   |||
Db 602 TGGTTACGTGAAATTTGTTATGTCGGATGTTGAACCGTGGTTCGCGTCTGATCCGTC 661

QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
   |||
Db 662 GTTCTGTGTCAACGTTAAGCGGCGGAGCTGCGGTTTACAAAATTTGGGTAAACCGCG 721

QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
   |||
Db 722 TTAGGACGCTATCTGTTACATCATCGACATTAAACCGGGACITTTATTGAGATAGCCGT 781

QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
   |||
Db 782 GATCGCGGGCTGTGGGGCGAGCTTCGCGCTCGGATTAAAGCGGTAACCGTGTGTGCTA 841

QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
   |||
Db 842 ACAGAACTGTTTACCGGCGTCACCGTGTGTAC 874

```

```

RESULT 8
ECOUBIA 2000 bp DNA linear BCT 26-APR-1993
LOCUS Escherichia coli chorismate lyase (ubiC), 4-hydroxybenzoate
DEFINITION octaprenyl transferase (ubiA) genes, complete cds, and
sn-glycerol-3-phosphate acyltransferase (plsB) genes, 3' end.
ACCESSION M93136
VERSION M93136.1 GI:148099
KEYWORDS 4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plsB
gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC
gene.
SOURCE Escherichia coli W3110
ORGANISM Escherichia coli W3110
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

Enterobacteriaceae; Escherichia.

1 (bases 1 to 2000)  
 Lightner, V.A., Bell, R.M. and Modrich, P.  
 The DNA sequences encoding plbB and dgk loci of Escherichia coli  
 J. Biol. Chem. 258 (18), 10856-10861 (1983)  
 6309817

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 PUBMED

2 (bases 1 to 2000)  
 Nichols, B.P. and Green, J.M.  
 Cloning and Sequencing of Escherichia coli ubiC and purification of  
 chorismate lyase  
 Unpublished (1992)  
 Original source text: Escherichia coli (sub\_strain W3110, strain  
 K-12) (library: Kohare lambda-1F8) DNA.  
 Location/Qualifiers  
 1..2000  
 /organism="Escherichia coli W3110"  
 /mol\_type="genomic DNA"  
 /strain="K-12"  
 /sub\_strain="W3110"  
 /db\_xref="taxon:316407"  
 /map="91.8 min"  
 /tissue\_lib="Kohare lambda-1F8"  
 380..877  
 /gene="ubiC"  
 380..877  
 /gene="ubiC"  
 /function="enzymatic chorismate --> p-hydroxybenzoate +  
 pyruvate"  
 /experiment="experimental evidence, no additional details  
 recorded"  
 /citation=[2]  
 /codon\_start=1  
 /transl\_table=11  
 /product="chorismate lyase"  
 /protein\_id="AAA24711.1"  
 /db\_xref="GI:148102"  
 /translation="MSHPALTLQRLALRYCKEIPALDPQLLDMLLEDLSMTKRFQOQK  
 TVSVIRREGVEQNEIPEELPLLPKESRYLREILLCADGEPMLAGTVVPVSTLSG  
 PELALQKLGKTPLGRLFTSSTLTDRDFEIGRDAGLWGRSRLSGKPLLTLELFLP  
 ASPLP"  
 890..1762  
 /gene="ubiA"  
 890..1762  
 /gene="ubiA"  
 /function="enzymatic - 3-octaprenyl-4-hydroxybenzoate  
 synthesis"  
 /experiment="experimental evidence, no additional details  
 recorded"  
 /citation=[2]  
 /codon\_start=1  
 /transl\_table=11  
 /product="4-hydroxybenzoate-octaprenyl transferase"  
 /protein\_id="AAA24712.1"  
 /db\_xref="GI:148101"  
 /translation="MNSLTQNKLAFLHRLMRTDKPIGALLLWFTLWALWVATPGVP  
 QLMILAVFAGVLMRAAGCVNDYADRFKDFGKVRKTRLPSPGAVTERARALFV  
 LVILFLLVLTNTWLTLLISAALAWYIPMKRYTHLPQWLGAAGSIPWAFAA  
 VSESVPLSCWLMFLANLWAVYDQYAMVDRDDVKIGIKSTAILFGYDKLLIGIL  
 QIGVLALMAIGELNGLGWYWSILVAGALFVYQOKLIANRERACPKAFMNNYVG  
 LVLFGLWSVWHF"  
 complement(1917..2000)  
 /gene="plbB"  
 complement(1917..2000)  
 /gene="plbB"  
 /EC\_number="2.3.1.15"  
 /function="lysophosphatidic acid synthesis"  
 /experiment="experimental evidence, no additional details  
 recorded"  
 /citation=[1]  
 /codon\_start=1  
 /transl\_table=11  
 /product="sn-glycerol-3-phosphate acyltransferase"  
 /protein\_id="AAA24713.1"

/db\_xref="GI:148102"  
 /translation="MKVVQLLAELITSDVRLTIESATQEG"

ORIGIN

Alignment Scores:  
 Pred. No.: 1.33e-78 Length: 2000  
 Score: 856.00 Matches: 166  
 Percent Similarity: 97.7% Conservatives: 1  
 Best Local Similarity: 97.1% Mismatches: 4  
 Query Match: 73.9% Indels: 0  
 DB: 15 Gaps: 0

US-10-718-311-8 (1-227) x ECOUBICA (1-2000)

Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76  
 Db 362 TGTAAAGAGAGATCGCGCATGTACACCCCGGTTAACGCAACTCGTGGCTGGCTAT 421

Qy 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96  
 Db 422 TGTAAAGAGATCCCTCGCTGATCCGCACTGCTCGACTGGCTGGCTGGAGATTCC 481

Qy 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
 Db 482 ATGACAAACCGTTTGTAAACACGACGAGGAAAAACGGTAAGCGTGACGATATCGCGAAGGG 541

Qy 117 PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136  
 Db 542 TTTGTGAGCAGATGAATCCCGAGNACTCGCGTGTCCGAAAGAGATCTGTTAC 601

Qy 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156  
 Db 602 TGGTTACGTGAATTTTGTATGTCCGATGGTGAACCGTGGCTGGCTGGTACCGCTC 661

Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176  
 Db 662 GTTCTGTGTCAACGTTAAAGCGGCGGAGTGGCGTTACAAAAATTGGGTAAACGCGG 721

Qy 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArg 196  
 Db 722 TTAGACGCTATCTGTTACATCATCGACATTACCCCGGACTTTATTGAGATAGCCGT 781

Qy 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216  
 Db 782 GATCGCGGCTGTGGGGCGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTA 841

Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227  
 Db 842 ACAGAACTGTTTTTACCAGCGGTACCGGTGTAC 874

RESULT 9  
 ECOUBIPLS 2000 bp DNA linear BCT 26-APR-1993  
 LOCUS Escherichia coli 4-hydroxybenzoate octaprenyl transferase (ubiA)  
 DEFINITION gene complete cds, chorismate lyase (ubiC) gene complete cds,  
 sn-glycerol-3-phosphate acyltransferase (plbB) gene, 3' end.  
 M93413  
 ACCESSION M93413.1 GI:148106  
 VERSION 4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plbB  
 KEYWORDS gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC  
 gene.

SOURCE Escherichia coli W3110  
 ORGANISM Escherichia coli W3110  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 2000)  
 AUTHORS Lightner, V.A., Bell, R.M. and Modrich, P.  
 TITLE The DNA sequences encoding plbB and dgk loci of Escherichia coli  
 JOURNAL J. Biol. Chem. 258 (18), 10856-10861 (1983)  
 PUBMED 6309817

REFERENCE 2 (bases 1 to 2000)  
 AUTHORS Nichols, B.P. and Green, J.M.  
 TITLE Cloning and sequencing of Escherichia coli ubiC and purification of  
 Cloning and sequencing of Escherichia coli ubiC and purification of

chorismate lyase  
J. Bacteriol. 174 (16), 5309-5316 (1992)  
1644758  
Original source text: Escherichia coli (sub\_strain W3110, strain K-12) (library: Kohara lambda) DNA.  
Location/Qualifiers  
1..2000  
/organism="Escherichia coli W3110"  
/mol\_type="genomic DNA"  
/strain="K-12"  
/sub\_strain="W3110"  
/db\_xref="taxon:316407"  
/map="91.8 min"  
/tissue\_lib="Kohara lambda"  
380..877  
/gene="ubiC"  
380..877  
/gene="ubiC"  
/function="enzymatic chorismate --> p-hydroxybenzoate + pyruvate"  
/experiment="experimental evidence, no additional details recorded"  
/note="N-terminal amino acid sequence of chorismate lyase -NH2-SHPALTLRALRYCEIPAL-"  
/citation=[2]  
/codon\_start=1  
/transl\_table=11  
/product="chorismate lyase"  
/protein\_id="AAA24716.1"  
/db\_xref="GI:148107"  
/translation="MSHPALTLRALRYCEIPALDPQLLDWLLLEDSTKRFQOQK TVSVTMIRGEFVEQNEPEELPLPKESRYWIREILLCADGEPFAGRTVVPVSLSG PELALQKLGKTPGLRYLTSTLTDRFIEIGRDAGLWGRSRRLSLGKPLLELFLP ASPLY"  
890..1762  
/gene="ubiA"  
890..1762  
/gene="ubiA"  
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate synthesis"  
/experiment="experimental evidence, no additional details recorded"  
/citation=[2]  
/number=2  
/codon\_start=1  
/transl\_table=11  
/product="4-hydroxybenzoate-octaprenyl transferase"  
/protein\_id="AAA24717.1"  
/db\_xref="GI:148108"  
/translation="MEWSLTQNKLLAFHRLMTDKPIGALLLLNPTLWALWVATPGVP QWLILAVFVAGVLMRAAGCVYNDYADRFKDVKRTANRPLPSGAVTEKEARALFV LVLSPLLVLTNTWTLISIALALAWYIPMKRYTHLPQVVLGAAGFWSIPMAFA VSESPLSCWLMFLANLWAVDTQYAMVDRDDVKIGIKSTAILFGYDKLIIGIL QIGVLALMAIIGELNGLGWYWSILVAGALFYQOKLIANRERACFRFNNVVG LVLFGLAMSYWHF"  
complement(1917..2000)  
/gene="plsB"  
complement(1917..2000)  
/gene="plsB"  
/EC\_number="2.3.1.15"  
/function="lysophosphatidic acid synthesis"  
/experiment="experimental evidence, no additional details recorded"  
/citation=[1]  
/codon\_start=1  
/transl\_table=11  
/product="sn-glycerol-3-phosphate acyltransferase"  
/protein\_id="AAA24718.1"  
/db\_xref="GI:148109"  
/translation="MKVYQLLAELITSDVRLTIESATQEGE"

gene  
CDS

gene  
CDS

gene  
CDS

ORIGIN  
Alignment Scores:

Pred. No.: 1,338-78 Length: 2000  
Score: 856.00 Matches: 166  
Percent Similarity: 97.7% Conservative: 1  
Best Local Similarity: 97.1% Mismatches: 4  
Query Match: 73.9% Indels: 0  
DB: 15 Gaps: 0  
US-10-718-311-8 (1-227) x ECOUBIPLS (1-2000)  
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76  
Db 362 TGTAAACGGAGAGTTCGGCATGTACACCCCGGTTAAACGCAACTGCGTGCCTGCGCTAT 421  
QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuAspSer 96  
Db 422 TGTAAAGAGATCCCTGCGCTGATCCGCAACTGCTGACTGCTGCTGCTGCTGCTGCT 481  
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
Db 482 ATGACAAAACGTTTGAACAGCAGGAGGAAAACGGTAAGCTGACGATGATCCGCGAGGG 541  
QY 117 PheValGluGlnAsnGluLeuProGluLeuProLeuLeuProLysGluSerArgTyr 136  
Db 542 TTTGTCGAGCAGATGAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 601  
QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156  
Db 602 TGGTTACGTGAAATTTTGTATGTCGCGATGCTGTAACCCGCTGCTGCTGCTGCTGCTG 661  
QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176  
Db 662 GTTCTGTGTCAACGTTTAAAGCGGCGGAGCTGGGCTTACAAAATTGGGTAAACGCCG 721  
QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeGlyArg 196  
Db 722 TTAGGACGCTATCTGTTACATCATCATCAATTAACCCGGGACTTTATTGAGATAGCCGT 781  
QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216  
Db 782 GATGCGCGGCTGTGGGGCGAGCTTCCCGCTGCGGATTAAGCGGTAAACCGCTGTTGCTA 841  
QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227  
Db 842 ACAGAACTGTTTTTACCGGCGTACCGTTGTAC 874  
RESULT 10  
ECUBIAC  
LOCUS  
DEFINITION  
E.coli ubiC and ubiA genes for chorismate lyase and 4-hydroxybenzoate octaprenyltransferase.  
X57434  
X57434.1 GI:43233  
4-hydroxybenzoate octaprenyl transferase; 4-hydroxybenzoate synthetase; chorismate lyase.  
Escherichia coli  
Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.  
REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
PUBLISHED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
Nishimura, K., Nakahigashi, K. and Inokuchi, H.  
Location of the ubiA gene on the physical map of Escherichia coli  
J. Bacteriol. 174 (17), 5762 (1992)  
1512213  
2 (bases 1 to 2034)  
Nishimura, K.  
Direct Submission  
Submitted (28-JAN-1991) K. Nishimura, Faculty of Sciences, Kyoto University, Sakyo-ku, Kyoto 606, Japan  
Location/Qualifiers  
1..2034  
/organism="Escherichia coli"  
/mol\_type="genomic DNA"  
/strain="K12 W3110"

```

-35_signal
-10_signal
gene
CDS

/db_xref="taxon:562"
/map="91.7 min"
/clone="634"
/clone_lib="Kohara library"
137..142
159..164
291..899
/gene="ubiC"
291..899
/gene="ubiC"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate synthetase"
/protein_id="CAA40681.1"
/db_xref="GI:43234"
/db_xref="GOA:P26602"
/db_xref="UniProtKB/Swiss-Prot:P26602"
/translacion="MRLRFCCVDHLCFTSPNFTLRYNFTLNGCFQMSHPALT
QRLARYCKEIPALDPQLLDLLEDSTKRFQOQKTVSMIREGFVEQNEIPEEL
PLLPKESRYWLREILLCADGEPWLAGRTVPVPSVLSGPELALQKLGKTLPLGRYLTSS
TLTRDFEIGRDAGLWGRSRLSLSGKPLLTLELFPASPLY"
912..1784
/gene="ubiA"
912..1784
/gene="ubiA"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate octaprenyltransferase"
/protein_id="CAA40682.1"
/db_xref="GI:43235"
/db_xref="GOA:P26601"
/db_xref="UniProtKB/Swiss-Prot:P26601"
/translacion="MWSLQNKLIHFRLMRTDKPIGALLLWPTLMALWVATPGVP
QLWTLAVAGVWLMRAAGCVNDYADRKFDGKHVRYKRTANRPLPSGAVTEKEARALFV
LVLISFLVLTNTWLTLLSAAALAWYPMKRYTHLPQVILGAAFGWSIPWAFAA
VSESVPLSCWLMFLANILWAVAYDTQYAMVDRDDDKIGIKSTAILFGQYDKLLIGIL
QIGVLAUMAIIGELNGWGYWSILVAGALFVYQKLIANREREAFCAPFMNNYVG
LVLFGLAMSWHF"

Alignment Scores:
Pred. No.: 1,368-78 Length: 2034
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0

US-10-718-311-8 (1-227) x ECUBIAC (1-2034)

Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db 384 TGTAACGGAGAGTTCGGCATGTGCACCCCGGTTAACGCAACTGCGTGGCTAT 443
Qy 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 444 TGTAAGAGATCCCTCGCTGATCCGCAACTGCTGCTGCTGCTGCTGCTGCTGCT 503
Qy 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 504 ATGACAAACGTTTTTGAACAGCAGGAGGAAACCGTAACGGTACGATGATTCGCGAAGG 563
Qy 117 PheValGluGlnAsnGluIleProGluLeuLeuProLeuProLeuProLysGluSerArgTyr 136
Db 564 TTTGTGCGAGAGATGAATCCCGAAGAACTGCGCTGCTGCTGCGGAAAGAGTCTGTTAC 623
Qy 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 624 TGGTTACGTGAATTTTGTATGTCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTC 683
Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 684 GTTCCTGTGTCAACGTTAAGCGGGCGGAGCTGGCGTTACAAAATTTGGGTAAACGCGG 743

```

```

Qy 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196
Db 744 TTAGAGCGTATCTGTTTCACATCATCGACATTAACCGGGACTTATTGAGATAGCCGT 803
Qy 197 AspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
Db 804 GATCCGGGCTGTGGGGCGACGTTCCCGCTCGCATTAAGCGGTAAACCGCTGTGTGCTA 863
Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 864 ACAGAACTGTTTTTACCGCGCTCACCGTTGTAC 896

RESULT 11
ECUBI
LOCUS ECUBI 2348 bp DNA linear BCT 09-SEP-2004
DEFINITION E.coli genes ubiC and ubiA.
ACCESSION X66619.1 S41687
VERSION X66619.1 GI:43230
KEYWORDS 4-hydroxybenzoate-octaprenyltransferase; chorismate-pyruvate lyase;
ubiA gene; ubiC gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 2348)
AUTHORS Siebert,M., Bechthold,A., Melzer,M., May,U., Berger,U.,
Schroder,G., Schroder,J., Severin,K. and Heide,L.
TITLE Ubiquinone biosynthesis. Cloning of the genes coding for chorismate
pyruvate-lyase and 4-hydroxybenzoate octaprenyl transferase from
Escherichia coli
JOURNAL FEBS Lett. 307 (3), 347-350 (1992)
PUBMED 1644192
REFERENCE 2 (bases 1 to 2348)
AUTHORS Heide,L.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1992) L. Heide, Inst of Pharmaceutical Biology,
Schaezlestrasse 1, 7800 Freiburg, FRG
COMMENT On Sep 24, 2004 this sequence version replaced gi:253104.
See also K00127, M93136 & M93413.

FEATURES
Location/Qualifiers
1..2348
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="MC4100"
/db_xref="taxon:562"
/map="91.8 min"
/clone="PALMUI"
/clone_lib="Sau3a partial digested DNA in pBluescript"
558..563
570..1067
/gene="ubiC"
570..1067
/gene="ubiC"
/codon_start=1
/transl_table=11
/product="chorismate-pyruvate lyase"
/protein_id="CAA47181.1"
/db_xref="GI:43231"
/db_xref="GOA:P26602"
/db_xref="UniProtKB/Swiss-Prot:P26602"
/translacion="MSPFALQRLALRYCKEIPALDPQLLDLLEDSTKRFQOQK
TVSTMIREGFVEQNEIPEELPLPKESRYWLREILLCADGEPWLAGRTVPVPSVLSG
PELALQKLGKTLPLGRYLTSSLTTRDFEIGRDAGLWGRSRLSLSGKPLLTLELFP
ASPLY"
1068..1073
1080..1952
/gene="ubiA"
1080..1952
/gene="ubiA"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"

```





U00096\_09 900001 1010000  
U00096\_10 1000001 1110000  
U00096\_11 1100001 1210000  
U00096\_12 1200001 1310000  
U00096\_13 1300001 1410000  
U00096\_14 1400001 1510000  
U00096\_15 1500001 1610000  
U00096\_16 1600001 1710000  
U00096\_17 1700001 1810000  
U00096\_18 1800001 1910000  
U00096\_19 1900001 2010000  
U00096\_20 2000001 2110000  
U00096\_21 2100001 2210000  
U00096\_22 2200001 2310000  
U00096\_23 2300001 2410000  
U00096\_24 2400001 2510000  
U00096\_25 2500001 2610000  
U00096\_26 2600001 2710000  
U00096\_27 2700001 2810000  
U00096\_28 2800001 2910000  
U00096\_29 2900001 3010000  
U00096\_30 3000001 3110000  
U00096\_31 3100001 3210000  
U00096\_32 3200001 3310000  
U00096\_33 3300001 3410000  
U00096\_34 3400001 3510000  
U00096\_35 3500001 3610000  
U00096\_36 3600001 3710000  
U00096\_37 3700001 3810000  
U00096\_38 3800001 3910000  
U00096\_39 3900001 4010000  
U00096\_40 4000001 4110000  
U00096\_41 4100001 4210000  
U00096\_42 4200001 4310000  
U00096\_43 4300001 4410000  
U00096\_44 4400001 4510000  
U00096\_45 4500001 4610000  
U00096\_46 4600001 4639675  
Continuation (43 of 47) of U00096 from base 4200001 (U00096 Escherichia coli K-12 MG1655)

Alignment Scores:  
Pred. No.: 1,77e-76 Length: 110000  
Score: 856.00 Matches: 166  
Percent Similarity: 97.7% Conservative: 1  
Best Local Similarity: 97.1% Mismatches: 4  
Query Match: 73.9% Indels: 0  
DB: 15 Gaps: 0  
US-10-718-311-8 (1-227) x U00096\_42 (1-110000)

Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76  
Db 50511 TGTAACGGAGAGTTCGGCATGTGCACCCCGGTAAACCACTGGCTGGCTAT 50570  
Qy 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96  
Db 50571 TGTAAGAGATCCCTCGCTGGATCCGCAACTGCTCGACTGGCTGGAGATTCC 50630  
Qy 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116  
Db 50631 ATGACAAACGTTTTGAACAGCAGGAGAAACGGTAACGGTACGATATCCGGAAGGG 50690  
Qy 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysLysSerArgTyr 136  
Db 50691 TTTGTGAGAGATGAATCCCGAAGACTGCCGCTGCTGCCAAGAGATCTCGTTAC 50750  
Qy 137 TrpLeuArgGluLeuLeuCysAlaAspGlyGluProTrpTrpLeuAlaGlyArgThrVal 156  
Db 50751 TGGTTACGTGAATTTTGTATGTGCCGATGGTAACCGTGGCTTGGCGGTCTACCGTC 50810  
Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176  
Db 50811 GTTCTGTGTCAAGTTAAGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAACCGCCG 50870

Qy 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluLeuGlyArg 196  
Db 50871 TTAGACGCTATCTGTTTCACATCATCGACATTAAACCCGGGACTTTATTGAGATAGCCGT 50930  
Qy 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216  
Db 50931 GATCCGGCTGTGGGGCGACGTTCCCGCTCGGATTAAACCGGTAAACCGCTGTGTGCTA 50990  
Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227  
Db 50991 ACAGAACTGTTTTTACCGCGCTCACCGTTGTAC 51023

RESULT 14  
AP009048\_42  
WPCOMMENT

Sequence split into 47 fragments LOCUS AP009048 Accession AP009048

Fragment Name Begin End

AP009048\_00 1 110000  
AP009048\_01 100001 210000  
AP009048\_02 200001 310000  
AP009048\_03 300001 410000  
AP009048\_04 400001 510000  
AP009048\_05 500001 610000  
AP009048\_06 600001 710000  
AP009048\_07 700001 810000  
AP009048\_08 800001 910000  
AP009048\_09 900001 1010000  
AP009048\_10 1000001 1110000  
AP009048\_11 1100001 1210000  
AP009048\_12 1200001 1310000  
AP009048\_13 1300001 1410000  
AP009048\_14 1400001 1510000  
AP009048\_15 1500001 1610000  
AP009048\_16 1600001 1710000  
AP009048\_17 1700001 1810000  
AP009048\_18 1800001 1910000  
AP009048\_19 1900001 2010000  
AP009048\_20 2000001 2110000  
AP009048\_21 2100001 2210000  
AP009048\_22 2200001 2310000  
AP009048\_23 2300001 2410000  
AP009048\_24 2400001 2510000  
AP009048\_25 2500001 2610000  
AP009048\_26 2600001 2710000  
AP009048\_27 2700001 2810000  
AP009048\_28 2800001 2910000  
AP009048\_29 2900001 3010000  
AP009048\_30 3000001 3110000  
AP009048\_31 3100001 3210000  
AP009048\_32 3200001 3310000  
AP009048\_33 3300001 3410000  
AP009048\_34 3400001 3510000  
AP009048\_35 3500001 3610000  
AP009048\_36 3600001 3710000  
AP009048\_37 3700001 3810000  
AP009048\_38 3800001 3910000  
AP009048\_39 3900001 4010000  
AP009048\_40 4000001 4110000  
AP009048\_41 4100001 4210000  
AP009048\_42 4200001 4310000  
AP009048\_43 4300001 4410000  
AP009048\_44 4400001 4510000  
AP009048\_45 4500001 4610000  
AP009048\_46 4600001 4646332

Continuation (43 of 47) of AP009048 from base 4200001 (AP009048 Escherichia coli W3110)

Alignment Scores:  
Pred. No.: 1,77e-76 Length: 110000  
Score: 856.00 Matches: 166  
Percent Similarity: 97.7% Conservative: 1  
Best Local Similarity: 97.1% Mismatches: 4  
Query Match: 73.9% Indels: 0



DB:	15	Gaps:	0	source
US-10-718-311-8 (1-227) x AP009048_42 (1-110000)				
Qy	57	CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr	76	/organism="Escherichia coli"
Db	56078	TGTAACGGAGAGTTCGGCATGTACACCCCGGTAAACGCACTGCGTGCCTAT	56137	/mol_type="unassigned DNA"
Qy	77	CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer	96	/strain="K-12"
Db	56138	TGTAAGAGATCCCTCCCTGGATCGCAACTGCTCGACTGGCTGTCTCTGGAGATTCC	56197	/sub_strain="MG1655"
Qy	97	MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly	116	/db_xref="taxon:562"
Db	56198	ATGACAAAACGTTTTGAACAGCAGGGAACCGTAAGCGTGACGATGATCCGGAAGGG	56257	/note="this sequence comprises the following lambda clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203, EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262, EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27; M13mp19 or Janus vectors were used for subcloning"
Qy	117	PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr	136	<1.1737
Db	56258	TTTGTCGAGCAGATGAATCCCGAAGAACTGCGCGTCTGCCGAAAGAGTCTCGTTAC	56317	/note="corresponds to M21516; ECOKATGA(1225. .2805)"
Qy	137	TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal	156	1..1257
Db	56318	TGTTACGTGAATTTTGTATGTCGCGATGGTGAACCGTGGCTTCCCGGTGCTACCGTC	56377	/gene="katG"
Qy	157	ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro	176	/gene="katG"
Db	56378	GTTCCTGTGTCAACGTTAAGCGGCGCGAGCTGGCGTTTCAAAAATTGGGTAAACCGCG	56437	/EC_number="1.11.1.6"
Qy	177	LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg	196	/standard_name="catalase HPI; catalase-peroxide; catalase"
Db	56438	TTAGGACGCTATCTGTTACATCATCATGACATTAACCCGGGACTTATTGAGATAGCCGT	56497	/note="CG Site No. 14983"
Qy	197	AspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu	216	/product="catalase hydroperoxidase I"
Db	56498	GATCCCGGCTGTGGGGCGAGCTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTA	56557	/protein_id="AAC43048.1"
Qy	217	ThrGluLeuPheLeuProAlaSerProLeuTyr	227	/db_xref="GI:396289"
Db	56558	ACAGAACTGTTTACCGCGGTACCGCTGTGATC	56590	/translation="AITSGLVVTQTPTQWSNYFEENLFKYEKVQTRSPAGAIQFEA LTHDMGPKSRVIGPEVKEDLIWQDLPQPIYNPTQDIIIDLKFAIDSGLSVSEVL SVYASASTFRGGDKRGAGARLALMPQRDWDVNAAVRALPVLEKIOKESGKSLA DIIVLVGVGVEKASAGLSIHVPFAPGRVDARQDQTDIEMFELLEPIADGPRNVRA RLDVSTTSLLIDKAQQLTLTAPENTALVGGMRVLGANGFGSKNGVFTDRVGLVSNDF RLNDMRYEYKATDESKELFEGRDRETGEVFTASRADLVFGSNSVLRAVAEYVASS DAHEKFXVDFFVAAVVKVMNLDKDFOLL"
RESULT 15				
ECOW89				
LOCUS	ECOW89 176195 bp DNA linear BCT 17-DEC-1993			
DEFINITION	E. coli chromosomal region from 89.2 to 92.8 minutes.			
ACCESSION	U00006			
VERSION	U00006.1 GI:409785			
KEYWORDS	Escherichia coli			
SOURCE	Escherichia coli			
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.			
REFERENCE	Blattner, F.R., Burland, V., Plunkett, G. III, Sofia, H.J. and Daniels, D.L.			
AUTHORS	Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes			
TITLE	Nucleic Acids Res. 21 (23), 5408-5417 (1993)			
JOURNAL	8265357			
PUBMED	2 (bases 1 to 176195)			
REFERENCE	Blattner, F.R.			
AUTHORS	Direct Submission			
TITLE	Submitted (03-SEP-1993) 608-263-7459			
JOURNAL	On Oct 29, 1993 this sequence version replaced gi:396288.			
COMMENT	This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps and conflicts with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW87 (119201) by 93 bp. NOTE: An update was submitted on 25-OCT-1993, reflecting a correction to the heme sequence.			
FEATURES				
Location/Qualifiers				

1..176195	/organism="Escherichia coli"
1..1257	/mol_type="unassigned DNA"
<1..1257	/strain="K-12"
/sub_strain="MG1655"	
/db_xref="taxon:562"	
/note="this sequence comprises the following lambda clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203, EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262, EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27; M13mp19 or Janus vectors were used for subcloning"	
<1..1737	
/note="corresponds to M21516; ECOKATGA(1225. .2805)"	
1..1257	/gene="katG"
<1..1257	
/gene="katG"	
/EC_number="1.11.1.6"	
/standard_name="catalase HPI; catalase-peroxide; catalase"	
/note="CG Site No. 14983"	
/codon_start=1	
/transl_table=11	
/product="catalase hydroperoxidase I"	
/protein_id="AAC43048.1"	
/db_xref="GI:396289"	
/translation="AITSGLVVTQTPTQWSNYFEENLFKYEKVQTRSPAGAIQFEA VDABEIIIDPDPKSKRPTMLVTDLTRDFPEFEKISRFLNDPQAFNFAFARXWPKF LTHDMGPKSRVIGPEVKEDLIWQDLPQPIYNPTQDIIIDLKFAIDSGLSVSEVL SVYASASTFRGGDKRGAGARLALMPQRDWDVNAAVRALPVLEKIOKESGKSLA DIIVLVGVGVEKASAGLSIHVPFAPGRVDARQDQTDIEMFELLEPIADGPRNVRA RLDVSTTSLLIDKAQQLTLTAPENTALVGGMRVLGANGFGSKNGVFTDRVGLVSNDF RLNDMRYEYKATDESKELFEGRDRETGEVFTASRADLVFGSNSVLRAVAEYVASS DAHEKFXVDFFVAAVVKVMNLDKDFOLL"	
<1..161	
/gene="katG"	
/note="corresponds to lambda clone EC18-126"	
1..93	
/gene="katG"	
/note="93 bp overlap with end of L19201 (ECOWU87)"	
156..15760	
/note="putative"	
1317..2255	
/note="similar to Desulfurolobus ambivalens hypoth. 28.3 kDa protein in sor 3' region"	
/codon_start=1	
/transl_table=11	
/label=ORF_0312	
/protein_id="AAC43049.1"	
/db_xref="GI:409786"	
/translation="MGERKLHTGLMSAAGKSNPLAISGLVLTIMWSYWIIFMKQVT SYIGAFQTFALRCIFGLAVLFVILLRGRMRPTPKYTLAIALLQTCGVAGKLAQVAL VSGGAGKVAIIISYTMPFWVIFALRIGRLRGQYFAILLAAFGFLVLPQWLOLQFS SMKSAMLAIIISGVSGASAIKAKLYARHPRVDLLSLTSWQMLYAAALVMSVALLVPO REIDWQPTVFWALAYSAILATALASLWLFVLLKNLPSLTSIASLSTLAVPVCGLVFSWML LGENPGAVEGGGIVLIVLALALVSRKKKEAVSVKRI"	
1429..1432	
/note="TTT in M21516; TTTT here"	
1720	
/note="C in M21516; G here"	
complement(2282..2899)	
/codon_start=1	
/transl_table=11	
/label=ORF_f205	
/protein_id="AAC43050.1"	
/db_xref="GI:396291"	
/translation="MKASLALLSLTFTSHLSKSPAVPTTVQIQTANTLAIADGAR QOIGSTLFPAYVQLTYPGDSVGVDRVIRALRSQKVDLQKLKHEDWAKNFA EYPOKWKLRDPSNIDHRRVPLETWSRHDKTRPTSKNPSDQAGDIVSRWLDNGLA	

```

HIGVSDGFARDGTPLVTHNIGAGAEEDVLFNRMVGHYRFVK"
complement(2886..2915)
/notes="promoter matrix score of 42; putative; within ORF
f205, which would suggest alternate start codon"
2912..3053
/notes="predicted bend of 81 degrees"
3031..3064
/standard_names="REP; repetitive extragenic palindromic
element"
/notes="contains 1 REP sequence"
3090..3219
/notes="predicted bend of 75 degrees"
complement(3174..4316)
/genes="glda"
complement(3174..4316)
/genes="glda"
/notes="similar to Bacillus stearothermophilus glycerol
dehydrogenase"
/translation="MPEHLLISKGAIMDRDIQSPGKIYQAGVINRLGEYKPLXER
WLVDGKFLVGFPAOSTVEKSPKAGLVVEIAPFGCECSQNEIDRLGIAETAQCGAIL
GIGGKTLDTAKALAHFMGVPAVAPTASTADPCSAHSVITDGEFDRVLLPNP
NMVIVDTXIVAGAPARLLAAGIGDALATWFEARACSRSGATMAGGKTOAALALAEI
CYNLTSEGERKMLAEORVTPALERVIEANTYLSRGVSGGLAAHVAHNGLITAI
PDAHVHGEKVAFGTQLVLNAPVEEIEETVAALSHAVGLPITLQAOLDIKEDVPAK
MRIVAAACAGSETIHNNPGGATDQVYVALLVADQYQRFQWE"
4022..4156
/notes="predicted bend of 75 degrees"
complement(4288..4950)
/notes="similar to Bacillus subtilis hypoth. 20 kDa
protein, in ter 3' region"
/codon_start=1
/translation="MELVLDITANVAVERLARIPPIAGVTNTPSIIAAKESIEVEVL
P
RLOKAIDEGILFAQWMSRDQGMVEAKRLRDAIPGIVKIPVTSGLAAIKLKE
GITTGTVAYAAQGLLAALAGAKYVAPVYNNRVDAGDGIPTVQELQTLLEHAPES
MVLAAFKTPQALDCLLAGCESITPLDVAQMLNTPAVESAIEKFEHDWNAFGTT
HL"
complement(4922..4949)
/notes="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/notes="similar to phosphotransferase system enzyme I"
/codon_start=1
/translation="MCSGAGGILTPISLLDNALGNLPAAGKVDABQSALENGLTIV
LKNTEFLRSDGATSLAEHRSLAGDTSIREHLLAGVAGLSCAEIVASANHCE
EFSRSSYLOERALDVRDVCFLLOQLIYGEORFPAPGKLTQPAICHADELTPSQFLE
LDKNHLGSLKSGTTSHTVILARSFNIPVLGVDIDALTPWQQQTIYIDNAGAIV
VEPEAVARYQQBARQDARLEQORVWLTOQARQADGIRIEIAANTAHSEVQAQAFG
NAGVGGLFRTEMLYMORTSAPGESLYNIQCALESANGRSIIVRTWIDIGDKFPDY
LNI PAEAMPFPYGRVRIYERYASLTFTQLRSILRASAHGSLKIMIPISSMEITLWY
KEKLAEAQQLRNHI PFDEKIQIGIMEVPSVMFIIDQCEETDFFSIGSNDITQYL
LAVDRNKAVTRHNSLNPAFLRALDYAVQVHRQKWI GLCGLGAKGSVLPILLVGL
GLDELSAPSIPAAKARMAQLDRECKRLNQAWACKTSLEVEHLLAQFRMTQODAP
LVTAECTLEDSWRSKEVLKGMTDNLNLLAGRCYPRKLEADRLWARLAVSTGLGFSF
AIPHSKSEHIFQSTISVARLOAPVRWGDDEAQFIIMLTLNKHAAGDQHMRIFSLRAR
IMHEFRNALVNAASADAIAISLLQHELEL"
7099..7428
/codon_start=1
/translation="
/label=ORF_o109

promoter
complement(4922..4949)
/notes="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/notes="similar to phosphotransferase system enzyme I"
/codon_start=1
/translation="MCSGAGGILTPISLLDNALGNLPAAGKVDABQSALENGLTIV
LKNTEFLRSDGATSLAEHRSLAGDTSIREHLLAGVAGLSCAEIVASANHCE
EFSRSSYLOERALDVRDVCFLLOQLIYGEORFPAPGKLTQPAICHADELTPSQFLE
LDKNHLGSLKSGTTSHTVILARSFNIPVLGVDIDALTPWQQQTIYIDNAGAIV
VEPEAVARYQQBARQDARLEQORVWLTOQARQADGIRIEIAANTAHSEVQAQAFG
NAGVGGLFRTEMLYMORTSAPGESLYNIQCALESANGRSIIVRTWIDIGDKFPDY
LNI PAEAMPFPYGRVRIYERYASLTFTQLRSILRASAHGSLKIMIPISSMEITLWY
KEKLAEAQQLRNHI PFDEKIQIGIMEVPSVMFIIDQCEETDFFSIGSNDITQYL
LAVDRNKAVTRHNSLNPAFLRALDYAVQVHRQKWI GLCGLGAKGSVLPILLVGL
GLDELSAPSIPAAKARMAQLDRECKRLNQAWACKTSLEVEHLLAQFRMTQODAP
LVTAECTLEDSWRSKEVLKGMTDNLNLLAGRCYPRKLEADRLWARLAVSTGLGFSF
AIPHSKSEHIFQSTISVARLOAPVRWGDDEAQFIIMLTLNKHAAGDQHMRIFSLRAR
IMHEFRNALVNAASADAIAISLLQHELEL"
7099..7428
/codon_start=1
/translation="
/label=ORF_o109

misc_structure
complement(7108..7136)
/notes="promoter matrix score of 46; putative"
7200..75839
/notes="corresponds to lambda clone EC21-52"
7578..7699
/notes="putative"
7771..8850
/notes="similar to phosphotransferase system enzyme II"
/codon_start=1
/translation="MPEHLLISKGAIMDRDIQSPGKIYQAGVINRLGEYKPLXER
WLVDGKFLVGFPAOSTVEKSPKAGLVVEIAPFGCECSQNEIDRLGIAETAQCGAIL
GIGGKTLDTAKALAHFMGVPAVAPTASTADPCSAHSVITDGEFDRVLLPNP
NMVIVDTXIVAGAPARLLAAGIGDALATWFEARACSRSGATMAGGKTOAALALAEI
CYNLTSEGERKMLAEORVTPALERVIEANTYLSRGVSGGLAAHVAHNGLITAI
PDAHVHGEKVAFGTQLVLNAPVEEIEETVAALSHAVGLPITLQAOLDIKEDVPAK
MRIVAAACAGSETIHNNPGGATDQVYVALLVADQYQRFQWE"
4022..4156
/notes="predicted bend of 75 degrees"
complement(4288..4950)
/notes="similar to Bacillus subtilis hypoth. 20 kDa
protein, in ter 3' region"
/codon_start=1
/translation="MELVLDITANVAVERLARIPPIAGVTNTPSIIAAKESIEVEVL
P
RLOKAIDEGILFAQWMSRDQGMVEAKRLRDAIPGIVKIPVTSGLAAIKLKE
GITTGTVAYAAQGLLAALAGAKYVAPVYNNRVDAGDGIPTVQELQTLLEHAPES
MVLAAFKTPQALDCLLAGCESITPLDVAQMLNTPAVESAIEKFEHDWNAFGTT
HL"
complement(4922..4949)
/notes="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/notes="similar to phosphotransferase system enzyme I"
/codon_start=1
/translation="MCSGAGGILTPISLLDNALGNLPAAGKVDABQSALENGLTIV
LKNTEFLRSDGATSLAEHRSLAGDTSIREHLLAGVAGLSCAEIVASANHCE
EFSRSSYLOERALDVRDVCFLLOQLIYGEORFPAPGKLTQPAICHADELTPSQFLE
LDKNHLGSLKSGTTSHTVILARSFNIPVLGVDIDALTPWQQQTIYIDNAGAIV
VEPEAVARYQQBARQDARLEQORVWLTOQARQADGIRIEIAANTAHSEVQAQAFG
NAGVGGLFRTEMLYMORTSAPGESLYNIQCALESANGRSIIVRTWIDIGDKFPDY
LNI PAEAMPFPYGRVRIYERYASLTFTQLRSILRASAHGSLKIMIPISSMEITLWY
KEKLAEAQQLRNHI PFDEKIQIGIMEVPSVMFIIDQCEETDFFSIGSNDITQYL
LAVDRNKAVTRHNSLNPAFLRALDYAVQVHRQKWI GLCGLGAKGSVLPILLVGL
GLDELSAPSIPAAKARMAQLDRECKRLNQAWACKTSLEVEHLLAQFRMTQODAP
LVTAECTLEDSWRSKEVLKGMTDNLNLLAGRCYPRKLEADRLWARLAVSTGLGFSF
AIPHSKSEHIFQSTISVARLOAPVRWGDDEAQFIIMLTLNKHAAGDQHMRIFSLRAR
IMHEFRNALVNAASADAIAISLLQHELEL"
7099..7428
/codon_start=1
/translation="
/label=ORF_o109

misc_feature
complement(7108..7136)
/notes="promoter matrix score of 46; putative"
7200..75839
/notes="corresponds to lambda clone EC21-52"
7578..7699
/notes="putative"
7771..8850
/notes="similar to phosphotransferase system enzyme II"
/codon_start=1
/translation="MPEHLLISKGAIMDRDIQSPGKIYQAGVINRLGEYKPLXER
WLVDGKFLVGFPAOSTVEKSPKAGLVVEIAPFGCECSQNEIDRLGIAETAQCGAIL
GIGGKTLDTAKALAHFMGVPAVAPTASTADPCSAHSVITDGEFDRVLLPNP
NMVIVDTXIVAGAPARLLAAGIGDALATWFEARACSRSGATMAGGKTOAALALAEI
CYNLTSEGERKMLAEORVTPALERVIEANTYLSRGVSGGLAAHVAHNGLITAI
PDAHVHGEKVAFGTQLVLNAPVEEIEETVAALSHAVGLPITLQAOLDIKEDVPAK
MRIVAAACAGSETIHNNPGGATDQVYVALLVADQYQRFQWE"
4022..4156
/notes="predicted bend of 75 degrees"
complement(4288..4950)
/notes="similar to Bacillus subtilis hypoth. 20 kDa
protein, in ter 3' region"
/codon_start=1
/translation="MELVLDITANVAVERLARIPPIAGVTNTPSIIAAKESIEVEVL
P
RLOKAIDEGILFAQWMSRDQGMVEAKRLRDAIPGIVKIPVTSGLAAIKLKE
GITTGTVAYAAQGLLAALAGAKYVAPVYNNRVDAGDGIPTVQELQTLLEHAPES
MVLAAFKTPQALDCLLAGCESITPLDVAQMLNTPAVESAIEKFEHDWNAFGTT
HL"
complement(4922..4949)
/notes="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/notes="similar to phosphotransferase system enzyme I"
/codon_start=1
/translation="MCSGAGGILTPISLLDNALGNLPAAGKVDABQSALENGLTIV
LKNTEFLRSDGATSLAEHRSLAGDTSIREHLLAGVAGLSCAEIVASANHCE
EFSRSSYLOERALDVRDVCFLLOQLIYGEORFPAPGKLTQPAICHADELTPSQFLE
LDKNHLGSLKSGTTSHTVILARSFNIPVLGVDIDALTPWQQQTIYIDNAGAIV
VEPEAVARYQQBARQDARLEQORVWLTOQARQADGIRIEIAANTAHSEVQAQAFG
NAGVGGLFRTEMLYMORTSAPGESLYNIQCALESANGRSIIVRTWIDIGDKFPDY
LNI PAEAMPFPYGRVRIYERYASLTFTQLRSILRASAHGSLKIMIPISSMEITLWY
KEKLAEAQQLRNHI PFDEKIQIGIMEVPSVMFIIDQCEETDFFSIGSNDITQYL
LAVDRNKAVTRHNSLNPAFLRALDYAVQVHRQKWI GLCGLGAKGSVLPILLVGL
GLDELSAPSIPAAKARMAQLDRECKRLNQAWACKTSLEVEHLLAQFRMTQODAP
LVTAECTLEDSWRSKEVLKGMTDNLNLLAGRCYPRKLEADRLWARLAVSTGLGFSF
AIPHSKSEHIFQSTISVARLOAPVRWGDDEAQFIIMLTLNKHAAGDQHMRIFSLRAR
IMHEFRNALVNAASADAIAISLLQHELEL"
7099..7428
/codon_start=1
/translation="
/label=ORF_o109

terminator
complement(7108..7136)
/notes="promoter matrix score of 46; putative"
7200..75839
/notes="corresponds to lambda clone EC21-52"
7578..7699
/notes="putative"
7771..8850
/notes="similar to phosphotransferase system enzyme II"
/codon_start=1
/translation="MPEHLLISKGAIMDRDIQSPGKIYQAGVINRLGEYKPLXER
WLVDGKFLVGFPAOSTVEKSPKAGLVVEIAPFGCECSQNEIDRLGIAETAQCGAIL
GIGGKTLDTAKALAHFMGVPAVAPTASTADPCSAHSVITDGEFDRVLLPNP
NMVIVDTXIVAGAPARLLAAGIGDALATWFEARACSRSGATMAGGKTOAALALAEI
CYNLTSEGERKMLAEORVTPALERVIEANTYLSRGVSGGLAAHVAHNGLITAI
PDAHVHGEKVAFGTQLVLNAPVEEIEETVAALSHAVGLPITLQAOLDIKEDVPAK
MRIVAAACAGSETIHNNPGGATDQVYVALLVADQYQRFQWE"
4022..4156
/notes="predicted bend of 75 degrees"
complement(4288..4950)
/notes="similar to Bacillus subtilis hypoth. 20 kDa
protein, in ter 3' region"
/codon_start=1
/translation="MELVLDITANVAVERLARIPPIAGVTNTPSIIAAKESIEVEVL
P
RLOKAIDEGILFAQWMSRDQGMVEAKRLRDAIPGIVKIPVTSGLAAIKLKE
GITTGTVAYAAQGLLAALAGAKYVAPVYNNRVDAGDGIPTVQELQTLLEHAPES
MVLAAFKTPQALDCLLAGCESITPLDVAQMLNTPAVESAIEKFEHDWNAFGTT
HL"
complement(4922..4949)
/notes="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
complement(4962..7097)
/notes="similar to phosphotransferase system enzyme I"
/codon_start=1
/translation="MCSGAGGILTPISLLDNALGNLPAAGKVDABQSALENGLTIV
LKNTEFLRSDGATSLAEHRSLAGDTSIREHLLAGVAGLSCAEIVASANHCE
EFSRSSYLOERALDVRDVCFLLOQLIYGEORFPAPGKLTQPAICHADELTPSQFLE
LDKNHLGSLKSGTTSHTVILARSFNIPVLGVDIDALTPWQQQTIYIDNAGAIV
VEPEAVARYQQBARQDARLEQORVWLTOQARQADGIRIEIAANTAHSEVQAQAFG
NAGVGGLFRTEMLYMORTSAPGESLYNIQCALESANGRSIIVRTWIDIGDKFPDY
LNI PAEAMPFPYGRVRIYERYASLTFTQLRSILRASAHGSLKIMIPISSMEITLWY
KEKLAEAQQLRNHI PFDEKIQIGIMEVPSVMFIIDQCEETDFFSIGSNDITQYL
LAVDRNKAVTRHNSLNPAFLRALDYAVQVHRQKWI GLCGLGAKGSVLPILLVGL
GLDELSAPSIPAAKARMAQLDRECKRLNQAWACKTSLEVEHLLAQFRMTQODAP
LVTAECTLEDSWRSKEVLKGMTDNLNLLAGRCYPRKLEADRLWARLAVSTGLGFSF
AIPHSKSEHIFQSTISVARLOAPVRWGDDEAQFIIMLTLNKHAAGDQHMRIFSLRAR
IMHEFRNALVNAASADAIAISLLQHELEL"
7099..7428
/codon_start=1
/translation="
/label=ORF_o109

CDS
US-10-718-311-8 (1-227) x ECOUW89 (1-176195)
Alignment Scores:
Pred. No.: 3.15e-76 Length: 176195
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0
US-10-718-311-8 (1-227) x ECOUW89 (1-176195)
Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db 117730 TGTAAACGAGAGATTTCGGCATGTTCACACCCCGGTAAACGCACTGCGTGCCTAT 117789
Qy 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 117790 TGTAAAGAGATCCCTGCCCTGGATCCGCACTGCTGCGCTGTTGCTGGAGAGATCC 117849
Qy 97 MetThrLysArgPheGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 116
Db 117850 ATGACAAAACGTTTTTGAACAGCAGGAGGAAAAACGTAAGCGTGACGATGATCCGCGAAGG 117909
Qy 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 117910 TTTGTCGAGCAGAGATGAATCCCGCAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 117969
Qy 137 TrpLeuArgGluLeuLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 117970 TGGTTACGTGAATTTTGTATTGTGCCGATGGTGAACCGTGGCTGCGGCTGCTACCGTC 118029
Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 118030 GTTCTGTGTCAACGTTTAAAGCGCGCGAGCTGCGCTTACAAAAATTTGGGTAAACGCG 118089
Qy 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196
Db 118090 TTAGGACGCTATCTGTTTCATCATTCACATTAACCCGGGACTTTTATTGAGATAGGCGGT 118149
Qy 197 AspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
Db 118150 GATGCCGGGCTGTGGGGGCGACGTTCCCGCTCGCATTAAGCGGTAACCGCTGTTGCTA 118209
Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 118210 ACAGAACTGTTTTTACCGGCGTCAACCGTTGTAT 118242
Search completed: July 11, 2006, 21:35:31
Job time : 5052.29 secs
```



**THIS PAGE BLANK (USPTO)**

November 2005

Published\_Applications\_Nucleic Acid and Published\_Applications\_Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 19:21:37 ; Search time 137 Seconds  
(without alignments)  
6491.174 Million cell updates/sec

Title: US-10-718-311-7

Perfect score: 684  
Sequence: 1 atgggtctctctgtctatttc.....cggcgtaaccgtgtactaa 684

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	144	21.1	762	8	US-11-247-587-125 Sequence 125, App
2	142.4	20.8	781	8	US-11-247-587-124 Sequence 124, App
3	124.6	18.2	698	8	US-11-247-587-128 Sequence 128, App
4	121.4	17.7	683	8	US-11-247-587-126 Sequence 126, App
5	119.8	17.5	700	8	US-11-247-587-127 Sequence 127, App
6	109.8	16.1	727	8	US-11-247-587-129 Sequence 129, App
7	85.2	12.5	718	6	US-10-953-349-10521 Sequence 10521, A
8	83.6	12.2	583	7	US-11-218-305-17631 Sequence 17631, A
9	83.6	12.2	658	7	US-11-218-305-17633 Sequence 17633, A
10	82	12.0	743	6	US-10-953-349-8894 Sequence 8894, App
11	80.4	11.8	640	7	US-11-218-305-17634 Sequence 17634, A
12	78.4	11.5	1023	8	US-11-216-545-4174 Sequence 4174, App
13	76.8	11.2	841	8	US-11-216-545-4173 Sequence 4173, App
14	76.2	11.1	1237	8	US-11-216-545-8381 Sequence 8381, App
15	74.4	10.9	799	6	US-10-953-349-9569 Sequence 9569, App
16	74.4	10.9	828	7	US-11-218-305-17632 Sequence 17632, A
17	46.2	6.8	627	8	US-11-216-545-6801 Sequence 6801, App
18	46.2	6.8	882	8	US-11-266-748A-36894 Sequence 36894, App
19	46.2	6.8	882	8	US-11-266-748A-388679 Sequence 388679, App
20	46.2	6.8	882	8	US-11-266-748A-450273 Sequence 450273, App
21	46.2	6.8	947	6	US-10-449-902-24132 Sequence 24132, A
22	39.8	5.8	845	6	US-10-449-902-19918 Sequence 19918, A
23	39.8	5.8	1939	6	US-10-449-902-17846 Sequence 17846, A
24	39.6	5.8	830	6	US-10-449-902-4174 Sequence 4174, App

25	38.2	5.6	2409	6	US-10-449-902-20231	Sequence 20231, A
26	37.8	5.5	899	6	US-10-449-902-18147	Sequence 18147, A
27	35.6	5.2	792	8	US-11-217-529-6134	Sequence 6134, App
28	35.6	5.2	1735	8	US-11-216-545-7862	Sequence 7862, App
29	33	4.8	1459	8	US-11-266-748A-381174	Sequence 381174, App
30	33	4.8	1459	8	US-11-266-748A-430930	Sequence 430930, App
31	32.8	4.8	1000	8	US-11-266-748A-222987	Sequence 222987, App
32	32.8	4.8	1000	8	US-11-266-748A-289901	Sequence 289901, App
33	32.8	4.8	1000	8	US-11-266-748A-292596	Sequence 292596, App
34	32.8	4.8	1000	8	US-11-266-748A-341330	Sequence 341330, App
35	32.8	4.8	1000	8	US-11-266-748A-344025	Sequence 344025, App
36	32.8	4.8	1000	8	US-11-266-748A-390405	Sequence 390405, App
37	32.8	4.8	1000	8	US-11-266-748A-401025	Sequence 401025, App
38	32.8	4.8	1000	8	US-11-266-748A-404362	Sequence 404362, App
39	32.8	4.8	1000	8	US-11-266-748A-404363	Sequence 404363, App
40	32.8	4.8	1000	8	US-11-266-748A-472071	Sequence 472071, App
41	32.8	4.8	1000	8	US-11-266-748A-475408	Sequence 475408, App
42	32.8	4.8	1000	8	US-11-266-748A-475409	Sequence 475409, App
43	32.8	4.8	1000	8	US-11-266-748A-481123	Sequence 481123, App
44	32.8	4.8	4378	8	US-11-266-748A-23487	Sequence 23487, A
45	32.8	4.8	4612	8	US-11-266-748A-24708	Sequence 24708, A

ALIGNMENTS

RESULT 1

US-11-247-587-125  
; Sequence 125, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 125  
; LENGTH: 762  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana  
US-11-247-587-125

Query Match 21.1%; Score 144; DB 8; Length 762;  
Best Local Similarity 86.4%; Pred. No. 8.7e-39;  
Matches 159; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY	1	ATGGCTTCCTCTGTCTCAATTTCTTCAGCAGCTGTTGCCACACGACGCAATGTTACCAAGCT	60
DB	27	ATGGCTTCCTCTGTCTCAATTTCTTCAGCAGCTGTTGCCACACGACGCAATGTTCTCAAGCT	86
QY	61	AGATGGTTGACCTTTCATCTGCTCTCAATTTCTTCAGCAGCTGTTGCCACGCAATGTTACCAAGCT	120
DB	87	AAATGGTTGACCTTTCATCTGCTCTCAATTTCTTCAGCAGCTGTTGCCACGCAATGTTCTCAAGCT	146
QY	121	CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTAGCTGCATGCAGGTG	180
DB	147	CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTAGCTGCATGCAGGTG	206
QY	181	TGGC 184	
DB	207	TGGC 210	

```
RESULT 2
US-11-247-587-124
; Sequence 124, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; FILE OF INVENTION: MicroRNAs
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 124
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-124
```

```
Query Match      20.8%; Score 142.4; DB 8; Length 781;
Best Local Similarity 85.9%; Pred. No. 3.1e-38;
Matches 158; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGGCTTCCTCTGTCAATTTCTTTCAGCAGCTTTGCGCACGCGAGCAATGTTACACAAGCT 60
Db 39 ATGGCTTCCTCTGTCAATTTCTTTCAGCAGCTTTGCGCACGCGAGCAATGTTGCTCAAGCT 98

Qy 61 AGCATGGTTGACCTTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 120
Db 99 AACATGGTTGACCTTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 158

Qy 121 CAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180
Db 159 CAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 218

Qy 181 TGGC 184
Db 219 TGGC 222
```

```
RESULT 3
US-11-247-587-128
; Sequence 128, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; FILE OF INVENTION: MicroRNAs
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
```

```
; SEQ ID NO 128
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-128

Query Match      18.2%; Score 124.6; DB 8; Length 698;
Best Local Similarity 82.9%; Pred. No. 3.7e-32;
Matches 155; Conservative 0; Mismatches 29; Indels 3; Gaps 1;

Qy 1 ATGGCTTCCTCTGTCAATTTCTTTCAGCAGCT---GTTGCGCACGCGAGCAATGTTACACAA 57
Db 20 ATGGCTTCCTCTGTGAATTTCTCTCAGCTGCTGCGTGGCACCGGCGCTAATGCTGCTCAA 79

Qy 58 GCTAGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAAGCACCTTTCCCTGTTTACAAAG 117
Db 80 GCCAGCATGGTTGCACCTTTCACTGGCCTCAAAATCTTCTCTCTCTCTCTCTCTCTCTCTCT 139

Qy 118 AAGCAAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCGAG 177
Db 140 AAGCAAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCGAG 199

Qy 178 GTGTGGC 184
Db 200 GTGTGGC 206
```

```
RESULT 4
US-11-247-587-126
; Sequence 126, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; FILE OF INVENTION: MicroRNAs
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 126
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-126
```

```
Query Match      17.7%; Score 121.4; DB 8; Length 683;
Best Local Similarity 81.8%; Pred. No. 4.6e-31;
Matches 153; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

Qy 1 ATGGCTTCCTCTGTCAATTTCTTTCAGCA---GCTTTGCGCACGCGAGCAATGTTACACAA 57
Db 37 ATGGCTTCCTCTAGTTATGTCCTCAGCAGCTGCTTTGCGACCGCGGCGCAATGCTGCTCAA 96

Qy 58 GCTAGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAAGCACCTTTCCCTGTTTACAAAG 117
Db 97 GCCAACATGGTTGACCTTCACTGGCTCAAGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 156

Qy 118 AAGCAAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCGAG 177
Db 157 AAGCAAAACCTTGACATCACTTCCATCTGCTAGCAATGTTGGAAGAGTTCAATGATGCGAG 216

Qy 178 GTGTGGC 184
Db 200 GTGTGGC 206
```



Db 217 GTGTGGC 223

## RESULT 5

US-11-247-587-127  
; Sequence 127, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3  
; SEQ ID NO 127  
; LENGTH: 700  
; TYPE: DNA  
; ORGANISM: Nicotiana benthamiana  
US-11-247-587-127

Query Match 17.5%; Score 119.8; DB 8; Length 700;

Best Local Similarity 81.3%; Pred. No. 1.7e-30;

Matches 152; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

QY 1 ATGGCTTCTCTGCTCATTTCTTCA---GCAGCTGTTGCCACACGCGCAAGCAATGTACACAA 57  
Db 2 ATGGCTTCTCTGCTCATTTCTTCA---GCAGCTGTTGCCACACGCGCAAGCAATGTCTCAA 61  
QY 58 GTAGCATGGTGGCAGCTTTCACCTGCTCAATCTTCAGCCACTTTCCCTGTACAAAG 117  
Db 62 GCCAGTATGTTGCACCTTTCACCTGCTCAAGTCCGCAACCTCTTCCCTGTTTCCAGA 121  
QY 118 AGCAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAG 177  
Db 122 AAACAAACCTTGACATCACTTCCATGCTAGCAACGCGGGAAGAGTTCAATGCATGCAG 181

QY 178 GTGTGGC 184

Db 182 GTGTGGC 188

## RESULT 6

US-11-247-587-129  
; Sequence 129, Application US/11247587  
; Publication No. US20060130176A1  
; GENERAL INFORMATION:  
; APPLICANT: The Rockefeller University  
; APPLICANT: Toboda, Jose Reyes  
; APPLICANT: Zhang, Xiuren  
; APPLICANT: Soyano, Takashi  
; APPLICANT: Chua, Nam-Hai  
; APPLICANT: Niu, Qi-Wen  
; APPLICANT: Lin, Shih-Shun  
; TITLE OF INVENTION: MicroRNAs  
; FILE REFERENCE: 2312-124  
; CURRENT APPLICATION NUMBER: US/11/247,587  
; CURRENT FILING DATE: 2005-10-12  
; PRIOR APPLICATION NUMBER: US 60/671,089  
; PRIOR FILING DATE: 2005-04-14  
; PRIOR APPLICATION NUMBER: PCT/US2004/033379  
; PRIOR FILING DATE: 2004-10-12  
; NUMBER OF SEQ ID NOS: 174

;; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3

; SEQ ID NO 129

; LENGTH: 727

; TYPE: DNA

; ORGANISM: Nicotiana benthamiana

US-11-247-587-129

## Query Match

Best Local Similarity 78.4%; Score 109.8; DB 8; Length 727;

Matches 145; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 3 GGCTTCTCTGCTCATTTCTTCA---GCAGCTGTTGCCACACGCGCAAGCAATGTACACAAGC 59

Db 7 GGCTTCTCTGCTCATTTCTTCA---GCAGCTGTTGCCACACGCGCAAGCAATGTCTTCAAGC 66

QY 60 TAGCATGGTGGCAGCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAA 119

Db 67 CAGCATGGTGGCAGCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAA 126

QY 120 GCAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGT 179

Db 127 ACAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGT 186

QY 180 GTGGC 184

Db 187 GTGGC 191

## RESULT 7

US-10-953-349-10521

; Sequence 10521, Application US/10953349

; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; FILE REFERENCE: ENCODED THERBY

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 10521

; LENGTH: 718

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-953-349-10521

## Query Match

Best Local Similarity 12.5%; Score 85.2; DB 6; Length 718;

Matches 130; Conservative 0; Mismatches 48; Indels 6; Gaps 1;

QY 1 ATGGCTTCTCTGCTCATTTCTTCAAGCAGCTGTTGCCACACGCGCAAGCAATGTACACAAGCT 60

Db 32 ATGGCTTCTCTGCTCATTTCTTCAAGCAGCTGTTGCCACACGCGCTGTTTAC-----CTCCCGGCTCAAGCC 85

QY 61 AGCATGGTGGCAGCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAAAG 120

Db 86 ACCATGGTGGCAGCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAAAG 145

QY 121 CAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

Db 146 GCCAACACGACATTAATCTTCCATCAACAGCAATGGGGGAGAGTTAGCTGCATGAAGGTG 205

QY 181 TGGC 184

Db 206 TGGC 209

## RESULT 8

US-11-218-305-17631

; Sequence 17631, Application US/11218305

; Publication No. US20060141495A1

; GENERAL INFORMATION:

```
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17631
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17631

Query Match      12.2%; Score 83.6; DB 7; Length 583;
Best Local Similarity 70.1%; Pred. No. 3.9e-18;
Matches 129; Conservative 0; Mismatches 49; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 60
   |||||
Db 33 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 86
   |||||

QY 61 AGCATGGTGGACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAG 120
   |||||
Db 87 ACCATGGTGGTCCATTACCGGCTTGAAGTCAATCCGCTGCAATCCCAAGTACCCGCAA 146
   |||||

QY 121 CAAAACCTTGACATCACTTCCATTTCTAGCAATGGTGAAGAGTGTAGCTGCATGCAAGTG 180
   |||||
Db 147 ACCAACACTGATATTACTTCCATTGCAAGCAAGGAGAGTGTAGCTGCATGCAAGTG 206
   |||||

QY 181 TGGC 184
   |||||
Db 207 TGGC 210
   |||||

RESULT 9
US-11-218-305-17633
; Sequence 17633, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17633
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17633

Query Match      12.2%; Score 83.6; DB 7; Length 658;
Best Local Similarity 70.1%; Pred. No. 4.1e-18;
Matches 129; Conservative 0; Mismatches 49; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 60
   |||||
Db 45 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 98
   |||||

QY 61 AGCATGGTGGACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAG 120
   |||||
```

```
Db 99 ACCATGGTGGCTCCATTCAACCGGCTTGAAGTCATCCGCTGCATTCCTCAGTCACCCGCAA 158
   |||||
QY 121 CAAAACCTTGACATCACTTCCATTTCTAGCAATGGTGAAGAGTGTAGCTGCATGCAAGTG 180
   |||||
Db 159 GCAAACCAACGACATTTACTTCCATTGCAAGCAACGAGGAGAGTGTAGCTGCATGAAGGTG 218
   |||||
QY 181 TGGC 184
   |||||
Db 219 TGGC 222
   |||||

RESULT 10
US-10-953-349-8894
; Sequence 8894, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8894
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8894

Query Match      12.0%; Score 82; DB 6; Length 743;
Best Local Similarity 69.6%; Pred. No. 1.6e-17;
Matches 128; Conservative 0; Mismatches 50; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 60
   |||||
Db 53 ATGGCTTCCTGTGCTATTTCTTACAGAGTGTGCGCAGCAGCAATGTTACACAAGCT 106
   |||||

QY 61 AGCATGGTGGACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAG 120
   |||||
Db 107 ACCATGGTGGCTCCATTACCGGCTTGAAGTCAATCCGCTGCAATCCCGGTACCCGCAAG 166
   |||||

QY 121 CAAAACCTTGACATCACTTCCATTTCTAGCAATGGTGAAGAGTGTAGCTGCATGCAAGTG 180
   |||||
Db 167 ACCAACCAAGGACATCACTTCCATCGCAAGCAACGCGGAGAGTGTAGCTGCATGAAGGTG 226
   |||||

QY 181 TGGC 184
   |||||
Db 227 TGGC 230
   |||||

RESULT 11
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634

Query Match      11.8%; Score 80.4; DB 7; Length 640;
Best Local Similarity 69.0%; Pred. No. 5.1e-17;
Matches 127; Conservative 0; Mismatches 51; Indels 6; Gaps 1;

QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCGAGCAATGTTACACAAGCT 60
   |||||
Db 33 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCGAGCAATGTTACACAAGCT 60
   |||||

QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
   |||||
Db 87 ACCATGGTGCCTCCATTACCGGCTTGAAGTCTTCCTCTGCATTCCAGTCCACCGCAAG 146
   |||||

QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGGTG 180
   |||||
Db 147 GCCAACACGACATTACCTCCATCGTTAGCAACGAGGAAGAGTTAGCTGCATGCAAGGTG 206
   |||||

QY 181 TGSC 184
   |||||
Db 207 TGSC 210
   |||||
```

```
RESULT 12
US-11-216-545-4174
; Sequence 4174, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4174
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-4174
```

```
Query Match      11.5%; Score 78.4; DB 8; Length 1023;
Best Local Similarity 71.7%; Pred. No. 3.2e-16;
Matches 132; Conservative 0; Mismatches 46; Indels 6; Gaps 2;

QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCGAGCAATGTTACACAAGCT 60
   |||||
Db 228 ATGGCTTCTCTCAATGATCTCTCCCGAGCTGTTACCACCTGTCACCGTG---CCGGTGCC 284
   |||||

QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
   |||||
Db 285 GGCATGGTGTCTCATCTCACTGCGCTCAAGTCCATGCTGGCTTCCCTCC---ACGAGGAAG 341
   |||||

QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGGTG 180
   |||||
Db 342 ACCAACAAATGACATTACCTCCATCTGCTAGCAACGGTGGAGAGTGAATGCAATGCAAGGTG 401
   |||||

QY 181 TGSC 184
   |||||
```

```
Db 402 TGSC 405
   |||||

RESULT 13
US-11-216-545-4173
; Sequence 4173, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4173
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-4173
```

```
Query Match      11.2%; Score 76.8; DB 8; Length 841;
Best Local Similarity 71.2%; Pred. No. 1e-15;
Matches 131; Conservative 0; Mismatches 47; Indels 6; Gaps 2;

QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACAGCGAGCAATGTTACACAAGCT 60
   |||||
Db 37 ATGGCTTCTCTCAATGATCTCTCCCGAGCTGTTACCAC---CGTCAACCGTGCCGGTGCC 93
   |||||

QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
   |||||
Db 94 GGCATGGTGTCTCCATTACCGGCTCAAAATCCATGGCTGGGCTTCCCTCC---ACGAGGAAG 150
   |||||

QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGGTG 180
   |||||
Db 151 ACCAACAAATGACATTACCTCCATGCTAGCAACGGTGGAGAGTGAATGCAATGCAAGGTG 210
   |||||

QY 181 TGSC 184
   |||||
Db 211 TGSC 214
   |||||
```

```
RESULT 14
US-11-216-545-8381
; Sequence 8381, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8381
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-8381
```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: July 11, 2006, 19:18:20 ; Search time 1466 Seconds  
(without alignments)  
5733.112 Million cell updates/sec  
Title: US-10-718-311-7  
Perfect score: 684  
Sequence: 1 aggcgttcctctgtcatttc.....cggcgtaccgtgtactaa 684  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 18892170 seqs, 6143817638 residues  
Total number of hits satisfying chosen parameters: 37784340  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description	Query Match	Length	DB ID	Description
1	684	100.0	684	3	US-09-855-341-7	Sequence 7, Appli			
2	684	100.0	684	3	US-09-896-866B-16	Sequence 16, Appl			
3	684	100.0	684	7	US-10-359-369-41	Sequence 41, Appl			
4	684	100.0	684	8	US-10-699-050-16	Sequence 16, Appl			
5	684	100.0	684	8	US-10-718-311-7	Sequence 7, Appli			
6	684	100.0	684	9	US-10-462-162-33	Sequence 33, Appl			
7	501	73.2	3452	9	US-10-462-162-47	Sequence 47, Appl			
8	499.4	73.0	1971	9	US-10-462-162-54	Sequence 54, Appl			
9	498	72.8	498	3	US-09-855-341-3	Sequence 29, Appli			
10	495	72.4	495	3	US-09-896-866B-12	Sequence 12, Appl			
11	495	72.4	495	7	US-10-359-369-37	Sequence 37, Appl			
12	495	72.4	495	8	US-10-699-050-12	Sequence 12, Appl			
13	495	72.4	495	8	US-10-718-311-3	Sequence 3, Appli			
14	495	72.4	495	10	US-10-450-763-26139	Sequence 26139, A			
15	472.8	69.1	1207	10	US-10-487-901-5765	Sequence 5765, Ap			
16	144	21.1	718	10	US-10-487-901-5765	Sequence 5765, Ap			
17	144	21.1	736	10	US-10-487-901-1704	Sequence 1704, Ap			

18	144	21.1	736	10	US-10-487-901-5043	Sequence 5043, Ap
19	142.4	20.8	297	10	US-10-487-901-7209	Sequence 7209, Ap
20	142.4	20.8	377	10	US-10-487-901-3504	Sequence 3504, Ap
21	142.4	20.8	489	10	US-10-487-901-3503	Sequence 3503, Ap
22	142.4	20.8	504	10	US-10-487-901-7207	Sequence 7207, Ap
23	140.8	20.6	204	3	US-09-839-477-3	Sequence 3, Appli
24	140.8	20.6	204	8	US-10-758-064-3	Sequence 3, Appli
25	137.6	20.1	668	10	US-10-487-901-7208	Sequence 7208, Ap
26	129.2	18.9	174	3	US-09-854-286-15	Sequence 15, Appl
27	128.8	18.8	169	6	US-10-165-420-3	Sequence 3, Appli
28	124.6	18.2	683	10	US-10-487-901-1713	Sequence 1713, Ap
29	124.6	18.2	737	10	US-10-487-901-5045	Sequence 5045, Ap
30	124.6	18.2	737	10	US-10-487-901-5766	Sequence 5766, Ap
31	121.4	17.7	738	10	US-10-487-901-1716	Sequence 1716, Ap
32	121.4	17.7	753	10	US-10-487-901-5764	Sequence 5764, Ap
33	121.4	17.7	754	10	US-10-487-901-5028	Sequence 5028, Ap
34	119.8	17.5	714	10	US-10-487-901-1703	Sequence 1703, Ap
35	119.8	17.5	736	10	US-10-487-901-5044	Sequence 5044, Ap
36	115	16.8	704	9	US-10-425-115-85509	Sequence 85509, A
37	113	16.5	151	7	US-10-321-434-4	Sequence 4, Appli
38	111.8	16.3	654	10	US-10-487-901-5042	Sequence 5042, Ap
39	111	16.2	180	10	US-10-487-901-88	Sequence 88, Appl
40	97.2	14.2	149	3	US-09-839-477-5	Sequence 5, Appli
41	97.2	14.2	149	8	US-10-758-064-5	Sequence 5, Appli
42	85.2	12.5	363	3	US-09-770-791-468	Sequence 468, App
43	85.2	12.5	711	3	US-09-910-664-32	Sequence 32, Appl
44	85.2	12.5	711	8	US-10-333-184-31	Sequence 31, Appl
45	85.2	12.5	739	8	US-10-333-184-182	Sequence 182, App

ALIGNMENTS

RESULT 1

US-09-855-341-7  
; Sequence 7, Application US/09855341  
; Patent No. US20020002715A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUIT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; TITLE OF INVENTION: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL  
US-09-855-341-7

Query Match	100.0%	Score 684;	DB 3;	Length 684;
Best Local Similarity	100.0%	Pred. No. 1.9e-222;		
Matches 684;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAGCT	60	
Db	1	ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAGCT	60	
QY	61	AGCATGGTTGACCTTTCCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG	120	
Db	61	AGCATGGTTGACCTTTCCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG	120	
QY	121	CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG	180	
Db	121	CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG	180	
QY	181	TGGCATATGTCACACCCCGCGTTACGCAACTTCGCTGCGCTATTGTTAAAGAGATC	240	

```

|||||
Db 181 TGGCATATGTACACCCCGCGTTAAACCAACTCGTGCCTGCGCTATTGTAAGAGATC 240
|||||
Qy 241 CCTGCCCTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAACGTT 300
|||||
Db 241 CCTGCCCTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAACGTT 300
|||||
Qy 301 TTTGAACAGCAGGGAAAAACGGTAAGCGTGACGATGATCCCGGAAGGGTTTGTGCGAGCAG 360
|||||
Db 301 TTTGAACAGCAGGGAAAAACGGTAAGCGTGACGATGATCCCGGAAGGGTTTGTGCGAGCAG 360
|||||
Qy 361 AATGAAATCCCCGAAGAACTGCCGCTCTGTCGCAAGAGTCTGTTACTGTTACGTGAA 420
|||||
Db 361 AATGAAATCCCCGAAGAACTGCCGCTCTGTCGCAAGAGTCTGTTACTGTTACGTGAA 420
|||||
Qy 421 ATTTGTTATGTGCCGATGTTGAACCGTGGCTTGC CGGTCGTACCGTCGTTCTGTGTCA 480
|||||
Db 421 ATTTGTTATGTGCCGATGTTGAACCGTGGCTTGC CGGTCGTACCGTCGTTCTGTGTCA 480
|||||
Qy 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
|||||
Db 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
|||||
Qy 541 CTGTTACATCATCGACATTAACCCGGGACTTTTATTCAGATAGCGCGTGATCCGGGCTG 600
|||||
Db 541 CTGTTACATCATCGACATTAACCCGGGACTTTTATTCAGATAGCGCGTGATCCGGGCTG 600
|||||
Qy 601 TGGGGGCGAGCTTCCCGCCTCGCANTTAAGCGGTAAACCGCTGTGCTAAACAGAACTGTTT 660
|||||
Db 601 TGGGGGCGAGCTTCCCGCCTCGCANTTAAGCGGTAAACCGCTGTGCTAAACAGAACTGTTT 660
|||||
Qy 661 TTACCGGCGTCACCGTTGTACTAA 684
|||||
Db 661 TTACCGGCGTCACCGTTGTACTAA 684
|||||

RESULT 2
US-09-896-866B-16
; Sequence 16, Application US/09896866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugates
; TITLE OF INVENTION: Benzoic Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; PRIORITY FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIORITY FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-09-896-866B-16

```

Qy	121	CAAAACCTTGACATCACTTCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG	181
Db	121	CAAAACCTTGACATCACTTCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG	180
Qy	181	TGGCATATGTACACACCCCGCTTAACGCAACTGGGTGCCTGCGCTATTGTAAAGAGATC	240
Db	181	TGGCATATGTACACACCCCGCTTAACGCAACTGGGTGCCTGCGCTATTGTAAAGAGATC	240
Qy	241	CTGCCCTTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATGACAAAACTG	300
Db	241	CTGCCCTTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATGACAAAACTG	300
Qy	301	TTTGAAACAGCAGGAGAAAAACGGTAAGCGGTGACGATGATCCCGGAAGGGTTGTCCGACGAG	360
Db	301	TTTGAAACAGCAGGAGAAAAACGGTAAGCGGTGACGATGATCCCGGAAGGGTTGTCCGACGAG	360
Qy	361	AATGAAATCCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTTGGTTACGTGAA	420
Db	361	AATGAAATCCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTTGGTTACGTGAA	420
Qy	421	ATTTTGTATTGTGCCGATGGTGAACCGTGGCTTGCCTGCTGTAACCGTGGTCTCTGTGTC	480
Db	421	ATTTTGTATTGTGCCGATGGTGAACCGTGGCTTGCCTGCTGTAACCGTGGTCTCTGTGTC	480
Qy	481	ACGTTAAGCGGGCCGGAGCTGGCGTTACAAAAATTTGGTAAAAACCGCGTTAGGACGCTAT	540
Db	481	ACGTTAAGCGGGCCGGAGCTGGCGTTACAAAAATTTGGTAAAAACCGCGTTAGGACGCTAT	540
Qy	541	CTGTTTCACATCATCGACATTAAACCGGGACATTATTGAGATAGGCGGTGATGCCGGGGCTG	600
Db	541	CTGTTTCACATCATCGACATTAAACCGGGACATTATTGAGATAGGCGGTGATGCCGGGGCTG	600
Qy	601	TGGGGGGCAGCTTCCCGCTCGCATTAAGCGGTAAACCGCTGTTGCTAACGAACTGTTT	660
Db	601	TGGGGGGCAGCTTCCCGCTCGCATTAAGCGGTAAACCGCTGTTGCTAACGAACTGTTT	660
Qy	661	TTACCGGGCTCACCGTTGTACTAA	684
Db	661	TTACCGGGCTCACCGTTGTACTAA	684

```

RESULT 3
US-10-359-369-41
; Sequence 41, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew
; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CL1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 41
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(684)
US-10-359-369-41

```

Db 1 ATGGCTTCCTCTGTCTATTTCTTCAGAGCTGTTCACACAGCAGCAATGTTACAAAGCT 60  
Qy 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCACTTTTCCCTGTTCACAAAGAAG 120  
Db 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCACTTTTCCCTGTTCACAAAGAAG 120  
Qy 121 CAATACTTGCACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGTGCATGCGAGGTG 180  
Db 121 CAATACTTGCACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGTGCATGCGAGGTG 180  
Qy 181 TGGCATATGTCACACCCCGGTTAAACCAACTGCGTGGCTGCGCTATTGTAAGAAGATC 240  
Db 181 TGGCATATGTCACACCCCGGTTAAACCAACTGCGTGGCTGCGCTATTGTAAGAAGATC 240  
Qy 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGGAGAGTTCCATGACAAACAGT 300  
Db 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGGAGAGTTCCATGACAAACAGT 300  
Qy 301 TTTGAACAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTGAGCAG 360  
Db 301 TTTGAACAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTGAGCAG 360  
Qy 361 AATGAATATCCCGAAGAACTGCGCTGCTCGAAAGAGTCTGTTACTGGTTACGTGAA 420  
Db 361 AATGAATATCCCGAAGAACTGCGCTGCTCGAAAGAGTCTGTTACTGGTTACGTGAA 420  
Qy 421 ATTTGTTATGTCGCGAGTGAACCGTGGCTGCTGCGGTGCTACCGTGGTTCTGTGTCA 480  
Db 421 ATTTGTTATGTCGCGAGTGAACCGTGGCTGCTGCGGTGCTACCGTGGTTCTGTGTCA 480  
Qy 481 ACCTTAAGCGGCGGAGCTGGCGTTACAAAATTTGGTTAAACCGCGTTAGACCGTAT 540  
Db 481 ACCTTAAGCGGCGGAGCTGGCGTTACAAAATTTGGTTAAACCGCGTTAGACCGTAT 540  
Qy 541 CTGTTACATCATCGACATTAACCCGGAGCTTTATTGAGATAGCGGTGATGCGCGCTG 600  
Db 541 CTGTTACATCATCGACATTAACCCGGAGCTTTATTGAGATAGCGGTGATGCGCGCTG 600  
Qy 601 TGGGGCGAGCTTCCCGCTCGCATTAAGCGGTAACCGGTGTTGCTTAACAGAACTGTTT 660  
Db 601 TGGGGCGAGCTTCCCGCTCGCATTAAGCGGTAACCGGTGTTGCTTAACAGAACTGTTT 660  
Qy 661 TTACCGGCGTCACCGTTGTACTAA 684  
Db 661 TTACCGGCGTCACCGTTGTACTAA 684

## RESULT 4

US-10-699-050-16

; Sequence 16, Application US/10699050

; Publication No. US20040142437A1

; GENERAL INFORMATION:

; APPLICANT: Flint, Dennis

; APPLICANT: Meyer, Knut

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugates

; FILE OF INVENTION: Benzoic Acid Glucosides

; FILE REFERENCE: BC1034 US NA

; CURRENT APPLICATION NUMBER: US/10/699,050

; CURRENT FILING DATE: 2003-10-30

; PRIOR APPLICATION NUMBER: US/09/896,866B

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16

; LENGTH: 684

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei  
US-10-699-050-16

Query Match 100.0%; Score 684; DB 8; Length 684;

Best Local Similarity 100.0%; Pred. No. 1.9e-222;

Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACAAAGCT 60  
Db 1 ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACAAAGCT 60  
Qy 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCACTTTTCCCTGTTCACAAAGAAG 120  
Db 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCACTTTTCCCTGTTCACAAAGAAG 120  
Qy 121 CAATACTTGCACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGTGCATGCGAGGTG 180  
Db 121 CAATACTTGCACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGTGCATGCGAGGTG 180  
Qy 181 TGGCATATGTCACACCCCGGTTAAACCAACTGCGTGGCTGCGCTATTGTAAGAAGATC 240  
Db 181 TGGCATATGTCACACCCCGGTTAAACCAACTGCGTGGCTGCGCTATTGTAAGAAGATC 240  
Qy 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGGAGAGTTCCATGACAAACAGT 300  
Db 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGGAGAGTTCCATGACAAACAGT 300  
Qy 301 TTTGAACAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTGAGCAG 360  
Db 301 TTTGAACAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTGAGCAG 360  
Qy 361 AATGAATATCCCGAAGAACTGCGCTGCTCGAAAGAGTCTGTTACTGGTTACGTGAA 420  
Db 361 AATGAATATCCCGAAGAACTGCGCTGCTCGAAAGAGTCTGTTACTGGTTACGTGAA 420  
Qy 421 ATTTGTTATGTCGCGAGTGAACCGTGGCTGCTGCGGTGCTACCGTGGTTCTGTGTCA 480  
Db 421 ATTTGTTATGTCGCGAGTGAACCGTGGCTGCTGCGGTGCTACCGTGGTTCTGTGTCA 480  
Qy 481 ACCTTAAGCGGCGGAGCTGGCGTTACAAAATTTGGTTAAACCGCGTTAGACCGTAT 540  
Db 481 ACCTTAAGCGGCGGAGCTGGCGTTACAAAATTTGGTTAAACCGCGTTAGACCGTAT 540  
Qy 541 CTGTTACATCATCGACATTAACCCGGAGCTTTATTGAGATAGCGGTGATGCGCGCTG 600  
Db 541 CTGTTACATCATCGACATTAACCCGGAGCTTTATTGAGATAGCGGTGATGCGCGCTG 600  
Qy 601 TGGGGCGAGCTTCCCGCTCGCATTAAGCGGTAACCGGTGTTGCTTAACAGAACTGTTT 660  
Db 601 TGGGGCGAGCTTCCCGCTCGCATTAAGCGGTAACCGGTGTTGCTTAACAGAACTGTTT 660  
Qy 661 TTACCGGCGTCACCGTTGTACTAA 684  
Db 661 TTACCGGCGTCACCGTTGTACTAA 684

## RESULT 5

US-10-718-311-7

; Sequence 7, Application US/10718311

; Publication No. US20040143867A1

; GENERAL INFORMATION:

; APPLICANT: Viitanen, Paul V.

; APPLICANT: MEYER, KNUT

; APPLICANT: VAN DYK, DREW

; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID

; FILE REFERENCE: BC1015 US NA

; CURRENT APPLICATION NUMBER: US/10/718,311

; CURRENT FILING DATE: 2003-11-20

; PRIOR APPLICATION NUMBER: US/09/855,341

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: MICROSOFT OFFICE 97

; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic CPL  
US-10-718-311-7

```
Query Match      100.0%; Score 684; DB 8; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT 60

Qy 61 AGCATGGTTGCACTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
Db 61 AGCATGGTTGCACTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

Qy 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

Qy 181 TGGCATATGTTCACACCCCGCGTTAAACGCAACTGCGTGGCGTCTGCTATTGTAAGAGATC 240
Db 181 TGGCATATGTTCACACCCCGCGTTAAACGCAACTGCGTGGCGTCTGCTATTGTAAGAGATC 240

Qy 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Db 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300

Qy 301 TTTGAACAGCAGGGAACGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Db 301 TTTGAACAGCAGGGAACGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360

Qy 361 AATGAAATCCCGGAAGAACTGCGCGTGTCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
Db 361 AATGAAATCCCGGAAGAACTGCGCGTGTCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420

Qy 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTGTGCGGTGTGACCGTCTCTCTGTGTCA 480
Db 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTGTGCGGTGTGACCGTCTCTCTGTGTCA 480

Qy 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540

Qy 541 CTGTTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600

Qy 601 TGGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660

Qy 661 TTACCGGCGTCACCGTTGTACTAA 684
Db 661 TTACCGGCGTCACCGTTGTACTAA 684
```

RESULT 6  
US-10-462-162-33  
; Sequence 33, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA  
; CURRENT APPLICATION NUMBER: US/10/462,162

; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 33  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-462-162-33

```
Query Match      100.0%; Score 684; DB 9; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT 60

Qy 61 AGCATGGTTGCACTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
Db 61 AGCATGGTTGCACTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

Qy 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

Qy 181 TGGCATATGTTCACACCCCGCGTTAAACGCAACTGCGTGGCGTCTGCTATTGTAAGAGATC 240
Db 181 TGGCATATGTTCACACCCCGCGTTAAACGCAACTGCGTGGCGTCTGCTATTGTAAGAGATC 240

Qy 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Db 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300

Qy 301 TTTGAACAGCAGGGAACGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Db 301 TTTGAACAGCAGGGAACGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360

Qy 361 AATGAAATCCCGGAAGAACTGCGCGTGTCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
Db 361 AATGAAATCCCGGAAGAACTGCGCGTGTCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420

Qy 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTGTGCGGTGTGACCGTCTCTCTGTGTCA 480
Db 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTGTGCGGTGTGACCGTCTCTCTGTGTCA 480

Qy 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540

Qy 541 CTGTTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600

Qy 601 TGGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660

Qy 661 TTACCGGCGTCACCGTTGTACTAA 684
Db 661 TTACCGGCGTCACCGTTGTACTAA 684
```

RESULT 7  
US-10-462-162-47  
; Sequence 47, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA



; CURRENT APPLICATION NUMBER: US/10/462,162  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 47  
; LENGTH: 3452  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence of nucleic acid fragment encoding CPL (from  
; OTHER INFORMATION: Escherichia coli), pHB 1-hydroxylase (from Candida  
; OTHER INFORMATION: parapsilosis), and UGT72B1 (from Arabidopsis thaliana).  
US-10-462-162-47

Query Match 73.2%; Score 501; DB 9; Length 3452;  
Best Local Similarity 100.0%; Pred. No. 1.8e-159;  
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 184 CATATGTACACCCCGGTTAAGCAACTGCGTGGCTGGCTGCTATTGTAAGAGATCCCT 243  
Db 1 CATATGTACACCCCGGTTAAGCAACTGCGTGGCTGGCTGCTATTGTAAGAGATCCCT 60  
QY 244 GCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTT 303  
Db 61 GCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTT 120  
QY 304 GAACAGCAGGGGAAAAACGGTAAGCGTACGATCCGCGAAGGGTTTTCGAGCAGAAT 363  
Db 121 GAACAGCAGGGGAAAAACGGTAAGCGTACGATCCGCGAAGGGTTTTCGAGCAGAAT 180  
QY 364 GAAATCCCGAAGAACTGCGTGGCTGCGGAAAGAGTCTGTTAGTGGTACGTGAAT 423  
Db 181 GAAATCCCGAAGAACTGCGTGGCTGCGGAAAGAGTCTGTTAGTGGTACGTGAAT 240  
QY 424 TTGTTATGTGCGATGCGTGAACCGTGGCTGCGGTAACCGTCTGCTGTTCAACG 483  
Db 241 TTGTTATGTGCGATGCGTGAACCGTGGCTGCGGTAACCGTCTGCTGTTCAACG 300  
QY 484 TTAAGCGGCGGAGCTGCGTTACAAAAATTGGGTAAACCGCTTAGACGCTATCTG 543  
Db 301 TTAAGCGGCGGAGCTGCGTTACAAAAATTGGGTAAACCGCTTAGACGCTATCTG 360  
QY 544 TTCACATCATCAGATTAACCGGACTTATTAGATAGGCGGTATCCGCGCTGTGG 603  
Db 361 TTCACATCATCAGATTAACCGGACTTATTAGATAGGCGGTATCCGCGCTGTGG 420  
QY 604 GGGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTA 663  
Db 421 GGGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTA 480  
QY 664 CCGCGCTACCGCTGTGTAATA 684  
Db 481 CCGCGCTACCGCTGTGTAATA 501

RESULT 8  
US-10-462-162-54  
; Sequence 54, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA  
; CURRENT APPLICATION NUMBER: US/10/462,162  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 54  
; LENGTH: 1971  
; TYPE: DNA

; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Nucleic acid sequence of a nucleic acid fragment inserted into an  
; OTHER INFORMATION: expression vector pET29a encoding CPL (from Escherichia coli) an  
; OTHER INFORMATION: pHB 1-hydroxylase (from Candida parapsilosis).  
US-10-462-162-54

Query Match 73.0%; Score 499.4; DB 9; Length 1971;  
Best Local Similarity 99.8%; Pred. No. 4.6e-159;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 184 CATATGTACACCCCGGTTAAGCAACTGCGTGGCTGGCTGCTATTGTAAGAGATCCCT 243  
Db 1 CATATGTACACCCCGGTTAAGCAACTGCGTGGCTGGCTGCTATTGTAAGAGATCCCT 60  
QY 244 GCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTT 303  
Db 61 GCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTT 120  
QY 304 GAACAGCAGGGGAAAAACGGTAAGCGTACGATCCGCGAAGGGTTTTCGAGCAGAAT 363  
Db 121 GAACAGCAGGGGAAAAACGGTAAGCGTACGATCCGCGAAGGGTTTTCGAGCAGAAT 180  
QY 364 GAAATCCCGAAGAACTGCGTGGCTGCGGAAAGAGTCTGTTAGTGGTACGTGAAT 423  
Db 181 GAAATCCCGAAGAACTGCGTGGCTGCGGAAAGAGTCTGTTAGTGGTACGTGAAT 240  
QY 424 TTGTTATGTGCGATGCGTGAACCGTGGCTGCGGTAACCGTCTGCTGTTCAACG 483  
Db 241 TTGTTATGTGCGATGCGTGAACCGTGGCTGCGGTAACCGTCTGCTGTTCAACG 300  
QY 484 TTAAGCGGCGGAGCTGCGTTACAAAAATTGGGTAAACCGCTTAGACGCTATCTG 543  
Db 301 TTAAGCGGCGGAGCTGCGTTACAAAAATTGGGTAAACCGCTTAGACGCTATCTG 360  
QY 544 TTCACATCATCAGATTAACCGGACTTATTAGATAGGCGGTATCCGCGCTGTGG 603  
Db 361 TTCACATCATCAGATTAACCGGACTTATTAGATAGGCGGTATCCGCGCTGTGG 420  
QY 604 GGGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTA 663  
Db 421 GGGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTA 480  
QY 664 CCGCGCTACCGCTGTGTAATA 684  
Db 481 CCGCGCTACCGCTGTGTAATA 501

RESULT 9  
US-10-462-162-29  
; Sequence 29, Application US/10462162  
; Publication No. US20040261147A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Flint, Dennis  
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes  
; FILE REFERENCE: CL 2155 US NA  
; CURRENT APPLICATION NUMBER: US/10/462,162  
; CURRENT FILING DATE: 2003-06-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 29  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-462-162-29

Query Match 72.8%; Score 498; DB 9; Length 498;  
Best Local Similarity 100.0%; Pred. No. 6.9e-159;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

187	Qy	ATGTCAACCCCGGTTAACGCAACTCGGTGGCTGGCGCTATTGTAAAGAGATCCCTCGCC	246
1	Db	ATGTCAACCCCGGTTAACGCAACTCGGTGGCTGGCGCTATTGTAAAGAGATCCCTCGCC	60
247	Qy	CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTTGAA	306
61	Db	CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTTGAA	120
307	Qy	CAGCAGGGAAAAACGTTAAGCGTAGCATGATCGCGAAGGGTTTGTGAGACAGATGA	366
121	Db	CAGCAGGGAAAAACGTTAAGCGTAGCATGATCGCGAAGGGTTTGTGAGACAGATGA	180
367	Qy	ATCCCCGAAGAACTGCCGCTGTCGCCGAAAGAGTCTCGTTACTGTGTTACGTGAAATTTTG	426
181	Db	ATCCCCGAAGAACTGCCGCTGTCGCCGAAAGAGTCTCGTTACTGTGTTACGTGAAATTTTG	240
427	Qy	TTATGTGCCGATGGTGAACCGGTGCTGCCGGTCGTACCGTCGTCTGTGTCAACGTTA	486
241	Db	TTATGTGCCGATGGTGAACCGGTGCTGCCGGTCGTACCGTCGTCTGTGTCAACGTTA	300
487	Qy	AGCGGCGCGAGCTGGCGTTACAAAAATTTGGTAAACCGCTTTAGACCGCTATCTGTTC	546
301	Db	AGCGGCGCGAGCTGGCGTTACAAAAATTTGGTAAACCGCTTTAGACCGCTATCTGTTC	360
547	Qy	ACATCATGACATTAAACCGGACCTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG	606
361	Db	ACATCATGACATTAAACCGGACCTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG	420
607	Qy	CGACGTTCCCGCTCGGATTAAACGGGTAAACCGCTGTTTGCTTAAACAGAACTGTTTTACCG	666
421	Db	CGACGTTCCCGCTCGGATTAAACGGGTAAACCGCTGTTTGCTTAAACAGAACTGTTTTACCG	480
667	Qy	GGCTACCGTTGTACTAA	684
481	Db	GGCTACCGTTGTACTAA	498

```

RESULT 10
US-09-855-341-3
; Sequence 3, Application US/09855341
; Patent No. US20020002715A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism.E. coli
US-09-855-341-3

```

[illegible]

```

RESULT 11
US-09-896-866B-12
; Sequence 12, Application US/09896866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate
; TITLE OF INVENTION: Benzoic Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896.866B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-896-866B-12

```

```
Query Match      72.4%; Score 495; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 7.3e-158;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy    187 ATGTCACACCCCGGTAAAGCAACTGGTGCCTATTGTAAGAAGATCCTGCC 246
      |||||
Db     1 ATGTCACACCCCGGTAAAGCAAATGGTGCGTGCCTATTGTAAGAAGATCCTGCC 60
```

Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTTCTGTGTCAACGTTA 300  
QY 487 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 546  
Db 301 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 360  
QY 547 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 606  
Db 361 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 420  
QY 607 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 666  
Db 421 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 480  
QY 667 GCGTACCGTTGTAC 681  
Db 481 GCGTACCGTTGTAC 495

## RESULT 12

US-10-359-369-37  
; Sequence 37, Application US/10359369  
; Publication No. US20030215927A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. duPont de Nemours and Company, Inc.  
; APPLICANT: Viitanen, Paul  
; APPLICANT: Meyer, Knut  
; APPLICANT: Van Dyk, Drew  
; TITLE OF INVENTION: UDP-Glucosyltransferases  
; FILE REFERENCE: CL1821 US NA  
; CURRENT APPLICATION NUMBER: US/10/359,369  
; CURRENT FILING DATE: 2003-02-06  
; PRIOR APPLICATION NUMBER: 60/355,511  
; PRIOR FILING DATE: 2002-02-07  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 37  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(495)  
US-10-359-369-37

Query Match 72.4%; Score 495; DB 7; Length 495;  
Best Local Similarity 100.0%; Pred. No. 7.3e-158;  
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 187 ATGTACACCCCGCTTAACGCACTGCGTGGCTGGCTATTGTAAAGAGATCCCTGCC 246  
Db 1 ATGTACACCCCGCTTAACGCACTGCGTGGCTGGCTATTGTAAAGAGATCCCTGCC 60  
QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 306  
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 120  
QY 307 CAGCAGGAAAAACGGTAAACGCTGATGATCCGCGAAGGGTTTCTCGAGCAGAAATGA 366  
Db 121 CAGCAGGAAAAACGGTAAACGCTGATGATCCGCGAAGGGTTTCTCGAGCAGAAATGA 180  
QY 367 ATCCCGGAAGAACTGCGGCTGCTGCCGAAGAGTCTCGTTACTGTTTACGTGAAATTTG 426  
Db 181 ATCCCGGAAGAACTGCGGCTGCTGCCGAAGAGTCTCGTTACTGTTTACGTGAAATTTG 240  
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTCAACGTTA 486  
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTCAACGTTA 300  
QY 487 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 546  
Db 301 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 360

QY 547 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 606  
Db 361 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 420  
QY 607 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 666  
Db 421 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 480  
QY 667 GCGTACCGTTGTAC 681  
Db 481 GCGTACCGTTGTAC 495

## RESULT 13

US-10-699-050-12  
; Sequence 12, Application US/10699050  
; Publication No. US20040142437A1  
; GENERAL INFORMATION:  
; APPLICANT: Flint, Dennis  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate  
; FILE REFERENCE: BC1034 US NA  
; CURRENT APPLICATION NUMBER: US/10/699,050  
; CURRENT FILING DATE: 2003-10-30  
; PRIOR APPLICATION NUMBER: US/09/896,866B  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/216,615  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-699-050-12

Query Match 72.4%; Score 495; DB 8; Length 495;  
Best Local Similarity 100.0%; Pred. No. 7.3e-158;  
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 187 ATGTACACCCCGCTTAACGCACTGCGTGGCTGGCTATTGTAAAGAGATCCCTGCC 246  
Db 1 ATGTACACCCCGCTTAACGCACTGCGTGGCTGGCTATTGTAAAGAGATCCCTGCC 60  
QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 306  
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 120  
QY 307 CAGCAGGAAAAACGGTAAACGCTGATGATCCGCGAAGGGTTTCTCGAGCAGAAATGA 366  
Db 121 CAGCAGGAAAAACGGTAAACGCTGATGATCCGCGAAGGGTTTCTCGAGCAGAAATGA 180  
QY 367 ATCCCGGAAGAACTGCGGCTGCTGCCGAAGAGTCTCGTTACTGTTTACGTGAAATTTG 426  
Db 181 ATCCCGGAAGAACTGCGGCTGCTGCCGAAGAGTCTCGTTACTGTTTACGTGAAATTTG 240  
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTCAACGTTA 486  
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTCAACGTTA 300  
QY 487 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 546  
Db 301 AGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGGTAGGACGCTATCTGTTTC 360  
QY 547 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 606  
Db 361 ACATCATCGACATTAACCGGGAATTTATTTAGATAGGCGGTGATCGCGGCTGTGGGG 420  
QY 607 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 666  
Db 421 CGACGTTCCCGCTGCGATTAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 480

Qy 667 GCGTACCGTGTGAC 681  
Db 481 GCGTACCGTGTGAC 495

## RESULT 14

US-10-718-311-3  
; Sequence 3, Application US/10718311  
; Publication No. US20040143867A1  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUZ  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; FILE REFERENCE: IN GREEN PLANTS  
; CURRENT APPLICATION NUMBER: US/10718,311  
; PRIOR FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: US/09/855,341  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 3  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: E. coli  
US-10-718-311-3

Query Match 72.4%; Score 495; DB 8; Length 495;  
Best Local Similarity 100.0%; Pred. No. 7.3e-158; Mismatches 0; Indels 0; Gaps 0;  
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 187 ATGTCACACCCCGGTTAACGCACTGCGTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 246  
Db 1 ATGTCACACCCCGGTTAACGCACTGCGTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 60  
Qy 247 CTGGATCCGCACTGCTCGACTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 306  
Db 61 CTGGATCCGCACTGCTCGACTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 120  
Qy 307 CAGCAGGGAACCGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGATGA 366  
Db 121 CAGCAGGGAACCGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGATGA 180  
Qy 367 ATCCCGGAAGAACTCCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTG 426  
Db 181 ATCCCGGAAGAACTCCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTG 240  
Qy 427 TTATGTCCGATGGTGAACCGTGGCTTCCGCTGCTACCGTTCCTGTCACAGTTA 486  
Db 241 TTATGTCCGATGGTGAACCGTGGCTTCCGCTGCTACCGTTCCTGTCACAGTTA 300  
Qy 487 AGCGGCGCGAGCTGGCTTACAAAATTGGTAAACCGCTTAGGACGCTATCTGTT 546  
Db 301 AGCGGCGCGAGCTGGCTTACAAAATTGGTAAACCGCTTAGGACGCTATCTGTT 360  
Qy 547 ACATCATCGAATTAAACCCGGAATTATTTAGATAGCGCGTGTGCGGGCTGTGGGG 606  
Db 361 ACATCATCGAATTAAACCCGGAATTATTTAGATAGCGCGTGTGCGGGCTGTGGGG 420  
Qy 607 CGAGCTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 666  
Db 421 CGAGCTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 480  
Qy 667 GCGTACCGTGTGAC 681  
Db 481 GCGTACCGTGTGAC 495

## RESULT 15

US-10-450-763-26139  
; Sequence 26139, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 26139  
; LENGTH: 1207  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (893)..(925)  
; OTHER INFORMATION: 81% homologous to Escherichia coli 4-hydroxybenzoate  
; OTHER INFORMATION: synthetase, accession number X57434, Smith-Waterman Score=49.  
US-10-450-763-26139

Query Match 69.1%; Score 472.8; DB 10; Length 1207;  
Best Local Similarity 99.2%; Pred. No. 4.5e-150;  
Matches 496; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
Qy 187 ATGTCACACCCCGGTTAACGCACTGCGTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 246  
Db 652 ATGTCACACCCCGGTTAACGCACTGCGTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 711  
Qy 247 CTGGATCCGCACTGCTCGACTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 306  
Db 712 CTGGATCCGCACTGCTCGACTGGCTGTGCTGAGGATTCATGACAAAACGTTTGA 771  
Qy 307 CAGCAGGGAACCGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGATGA 366  
Db 772 CAGCAGGGAACCGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGATGA 831  
Qy 367 ATCCCGGAAGAACTCCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTG 426  
Db 832 ATCCCGGAAGAACTCCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTG 891  
Qy 427 T-TATGTCCGATGGTGAACCGTGGCTTGGC - GGTGCTACCGTCTGTTCTCTGTCAACGT 484  
Db 892 TGTATGTCCGATGGCGAACC GGGCTTGGCGGCTCGTACCGTCTGTTCTGTCAACGT 951  
Qy 485 TAACCGGCGCGAGCTGGCTTACAAAATTGGGTAAACCGCTTAGGACGCTATCTGT 544  
Db 952 TAACCGGCGCGAGCTGGCTTACAAAATTGGGTAAACCGCTTAGGACGCTATCTGT 1011  
Qy 545 TCACATCATCGAATTAAACCGGACCTTTATTAGATAGAGCGGTGATCCCGGGCTGTGGG 604  
Db 1012 TCACATCATCGAATTAAACCGGACCTTTATTAGATAGAGCGGTGATCCCGGGCTGTGGG 1071  
Qy 605 GCGACGTTCCCGCTCGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTAC 664  
Db 1072 GCGACGTTCCCGCTCGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTAC 1131  
Qy 565 GCGCTCACCGTGTACTAA 684  
Db 1132 GCGCTCACCGTGTACTAA 1151

Search completed: July 11, 2006, 19:46:02  
Job time : 1467 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 19:14:52 ; Search time 190 Seconds  
(without alignments)  
6735.992 Million cell updates/sec

Title: US-10-718-311-7  
Perfect score: 684  
Sequence: 1 acggcttcctgtcatttc.....cggcgaccgttgtactaa 684

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq: \*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq: \*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq: \*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq: \*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq: \*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq: \*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq: \*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq: \*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq: \*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	684	100.0	684	3	US-09-896-866B-16
2	684	100.0	684	3	US-09-855-341-7
3	498	72.8	854	3	US-09-064-693A-24
4	498	72.8	6641	3	US-09-064-693A-25
5	495	72.4	495	3	US-09-896-866B-12
6	495	72.4	495	3	US-09-855-341-3
7	282	41.2	645	3	US-09-489-039A-1234
8	140.8	20.6	204	3	US-09-839-477-3
9	140.8	20.6	499	3	US-09-495-797-41
10	133.2	19.5	1442	2	US-08-152-483B-8
11	130.6	19.1	177	2	US-08-095-726-12
12	130.6	19.1	177	2	US-08-096-043-12
13	130.6	19.1	177	2	US-08-093-577-8
14	130.6	19.1	177	2	US-08-331-004A-6
15	130.6	19.1	177	2	US-08-096-623A-12
16	130.6	19.1	177	2	PCT-US95-13937A-6
17	126.6	18.5	537	3	US-09-543-681A-3122
18	77.2	14.2	149	3	US-09-839-477-5
19	77.2	11.3	281	2	US-08-469-202-15
20	77.2	11.3	281	2	US-08-469-202-17
21	77.2	11.3	281	2	US-08-469-202-19
22	77.2	11.3	281	2	US-08-484-434C-15
23	77.2	11.3	281	2	US-08-484-434C-18

24	77.2	11.3	281	2	US-08-484-434C-22	Sequence 22, Appl
25	77.2	11.3	281	3	US-09-384-361-15	Sequence 15, Appl
26	77.2	11.3	281	3	US-09-384-361-18	Sequence 18, Appl
27	77.2	11.3	281	3	US-09-384-361-22	Sequence 22, Appl
28	75.2	11.0	104	2	US-08-095-726-38	Sequence 38, Appl
29	75.2	11.0	104	2	US-08-096-043-35	Sequence 35, Appl
30	75.2	11.0	104	2	US-08-093-577-31	Sequence 31, Appl
31	75.2	11.0	104	2	US-08-096-623A-43	Sequence 43, Appl
32	74.4	10.9	264	3	US-09-441-340-9	Sequence 9, Appl
33	74.4	10.9	268	3	US-09-186-002-5	Sequence 5, Appl
34	74.4	10.9	279	2	US-08-391-339-9	Sequence 9, Appl
35	74.4	10.9	279	2	US-08-484-274A-9	Sequence 9, Appl
36	74.4	10.9	279	2	US-09-612-404-9	Sequence 9, Appl
37	74.4	10.9	355	2	US-08-090-523-5	Sequence 5, Appl
38	74.4	10.9	355	2	US-08-398-627-5	Sequence 5, Appl
39	74.4	10.9	355	2	US-08-406-858-5	Sequence 5, Appl
40	74.4	10.9	355	2	US-08-476-519-7	Sequence 7, Appl
41	74.4	10.9	355	3	US-08-120-703A-5	Sequence 5, Appl
42	74.4	10.9	355	3	US-08-399-023-5	Sequence 5, Appl
43	74.4	10.9	355	7	PCT-US91-04036-5	Sequence 5, Appl
44	74.4	10.9	355	7	PCT-US94-05275-5	Sequence 5, Appl
45	74.4	10.9	355	7	PCT-US95-09323-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-09-896-866B-16  
; Sequence 16, Application US/0989686B  
; Patent No. 6642036  
; GENERAL INFORMATION:  
; APPLICANT: Flint, Dennis  
; APPLICANT: Meyer, Knut  
; APPLICANT: Viitanen, Paul  
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate  
; FILE REFERENCE: BC1034 US NA  
; CURRENT APPLICATION NUMBER: US/09/896, 866B  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/216, 615  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; TYPE: DNA  
; ORGATISM: Artificial Sequence  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei  
US-09-896-866B-16

Query Match	100.0%;	Score 684;	DB 3;	Length 684;
Best Local Similarity	100.0%;	Pred. No. 1.1e-230;		
Matches	684;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGGCTTCCTGTGTCATTTCCTTTCAGCAGCTGTTCACGAGCAATGTTACCAAGCT	60	
Db	1	ATGGCTTCCTGTGTCATTTCCTTTCAGCAGCTGTTCACGAGCAATGTTACCAAGCT	60	
QY	61	ACATGGTTGCACCTTTCACCTGCTCTCAATTCCTAGCCACTTTCCTCTTCAAGAGAAG	120	
Db	61	ACATGGTTGCACCTTTCACCTGCTCTCAATTCCTAGCCACTTTCCTCTTCAAGAGAAG	120	
QY	121	CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG	180	
Db	121	CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG	180	
QY	181	TGGCATATGTCAACCCCGGTTTAACGAACTGCGTGCCTATTGTTAAAGAGATC	240	
Db	181	TGGCATATGTCAACCCCGGTTTAACGAACTGCGTGCCTATTGTTAAAGAGATC	240	

QY 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300  
DB 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300  
QY 301 TTGTAACAGCAGGAGGAAAAACGGTAAGCGTGCAGATGATCCGCGAAGGGTTTGTGCGAGCAG 360  
DB 301 TTGTAACAGCAGGAGGAAAAACGGTAAGCGTGCAGATGATCCGCGAAGGGTTTGTGCGAGCAG 360  
QY 361 AATGAATCCCGGAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTAGCTGAA 420  
DB 361 AATGAATCCCGGAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTAGCTGAA 420  
QY 421 ATTTTGTATGTGCGCATGTTGAAACCGTGGCTTTCGCGGTGCTGACCGTCTGCTCTGTGCA 480  
DB 421 ATTTTGTATGTGCGCATGTTGAAACCGTGGCTTTCGCGGTGCTGACCGTCTGCTCTGTGCA 480  
QY 481 AGCTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540  
DB 481 AGCTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540  
QY 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600  
DB 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600  
QY 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660  
DB 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660  
QY 661 TTACCGGGCTACCGTTGTACTAA 684  
DB 661 TTACCGGGCTACCGTTGTACTAA 684

## RESULT 2

US-09-855-341-7  
; Sequence 7, Application US/09855341  
; Patent No. 6683231  
; GENERAL INFORMATION:  
; APPLICANT: VIITANEN, PAUL V.  
; APPLICANT: MEYER, KNUT  
; APPLICANT: VAN DYK, DREW  
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID  
; FILE REFERENCE: IN GREEN PLANTS  
; FILE REFERENCE: BC1015 US NA  
; CURRENT APPLICATION NUMBER: US/09/855,341  
; CURRENT FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 7  
; LENGTH: 684  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL  
US-09-855-341-7

Query Match 100.0%; Score 684; DB 3; Length 684;  
Best Local Similarity 100.0%; Pred. No. 1.1e-230;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCTCTGTGCTATTTCTTACAGAGCTGTTGTCACACGAGCAATGTTACACAAGCT 60  
DB 1 ATGGCTTCTCTGTGCTATTTCTTACAGAGCTGTTGTCACACGAGCAATGTTACACAAGCT 60  
QY 61 AGCATGGTTGCACCTTTCACCTGCTCAAACTTCAGCACTTTCCTGTTTACAAAGAG 120  
DB 61 AGCATGGTTGCACCTTTCACCTGCTCAAACTTCAGCACTTTCCTGTTTACAAAGAG 120  
QY 121 CAAACCTTGACATCATTCTTACCTAGCAATGTTGGAAGAGTTAGCTGATGAGGTG 180  
DB 121 CAAACCTTGACATCATTCTTACCTAGCAATGTTGGAAGAGTTAGCTGATGAGGTG 180  
QY 181 TGGCATATGTACACCCCGGCTTAACGCAACTGCGGTGCGCTATTGTAAGAGATC 240

DB 181 TGGCATATGTACACCCCGGCTTAACGCAACTGCGGTGCGCTATTGTAAGAGATC 240  
QY 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300  
DB 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300  
QY 301 TTGTAACAGCAGGAGGAAAAACGGTAAGCGTGCAGATGATCCGCGAAGGGTTTGTGCGAGCAG 360  
DB 301 TTGTAACAGCAGGAGGAAAAACGGTAAGCGTGCAGATGATCCGCGAAGGGTTTGTGCGAGCAG 360  
QY 361 AATGAATCCCGGAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTAGCTGAA 420  
DB 361 AATGAATCCCGGAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTAGCTGAA 420  
QY 421 ATTTTGTATGTGCGCATGTTGAAACCGTGGCTTTCGCGGTGCTGACCGTCTGCTCTGTGCA 480  
DB 421 ATTTTGTATGTGCGCATGTTGAAACCGTGGCTTTCGCGGTGCTGACCGTCTGCTCTGTGCA 480  
QY 481 ACCTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540  
DB 481 ACCTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540  
QY 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600  
DB 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600  
QY 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660  
DB 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660  
QY 661 TTACCGGGCTACCGTTGTACTAA 684  
DB 661 TTACCGGGCTACCGTTGTACTAA 684

## RESULT 3

US-09-064-693A-24  
; Sequence 24, Application US/09064693A  
; Patent No. 6210937  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Thomas E.  
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY  
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION  
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: W. Gary Goodson  
; ADDRESSEE: INEEL--Lockheed Martin Idaho  
; ADDRESSEE: Technologies Co.  
; STREET: P.O. Box 1625  
; CITY: Idaho Falls  
; STATE: Idaho  
; COUNTRY: USA  
; ZIP: 83415-3810  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
; COMPUTER: Toshiba Satellite Pro T2150CDS  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: Word Perfect 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/064,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: W. Gary Goodson  
; REGISTRATION NUMBER: 22,387  
; REFERENCE/DOCKET NUMBER: LIT-PI-296  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (208)526-9469

TELEFAX: (208)526-8339  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 854 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-064-693A-24

Query Match 72.8%; Score 498; DB 3; Length 854;  
Best Local Similarity 100.0%; Pred. No. 6.2e-165;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 187 ATGTACACCCCGGTTAAGCAACTGCGTGGCGCTATTGTTAAAGAGATCCCTGCC 246  
Db 321 ATGTACACCCCGGTTAAGCAACTGCGTGGCGCTATTGTTAAAGAGATCCCTGCC 380  
QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 306  
Db 381 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 440  
QY 307 CAGCAGGGGAAAAACGGTAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGAA 366  
Db 441 CAGCAGGGGAAAAACGGTAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGAA 500  
QY 367 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAAATTTTG 426  
Db 501 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAAATTTTG 560  
QY 427 TTATGTGCGGATGGTGAACCGTGGCTTGCCTGTTACCGTCTGTTTACCG 486  
Db 561 TTATGTGCGGATGGTGAACCGTGGCTTGCCTGTTACCGTCTGTTTACCG 620  
QY 487 AGCGGCGCGGAGTGGCGTTTACAAAATTTGGGTAAACCGCGTTAGGAGCGTATCTGTTTC 546  
Db 621 AGCGGCGCGGAGTGGCGTTTACAAAATTTGGGTAAACCGCGTTAGGAGCGTATCTGTTTC 680  
QY 547 ACATCATGACATTAAACCGGAGCTTTATTTAGATAGGCGGCTGATCCGGGCTGTGGGG 606  
Db 681 ACATCATGACATTAAACCGGAGCTTTATTTAGATAGGCGGCTGATCCGGGCTGTGGGG 740  
QY 607 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTACCG 666  
Db 741 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTACCG 800  
QY 667 GCGTCACCGTTGTACTAA 684  
Db 801 GCGTCACCGTTGTACTAA 818

RESULT 4  
US-09-064-693A-25  
Sequence 25, Application US/09064693A  
Patent No. 6210937  
GENERAL INFORMATION:  
APPLICANT: Ward, Thomas E.  
TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY  
ENGINEERED BACTERIA FOR PRODUCTION  
OF A SPECIFIC PLASTICS PRECURSOR  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: W. Gary Goodson  
ADDRESS: INEL--lockheed Martin Idaho  
ADDRESSEE: Technologies Co.  
STREET: P.O. Box 1625  
CITY: Idaho Falls  
STATE: Idaho  
COUNTRY: USA  
ZIP: 83415-3810  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: Toshiba Satellite Pro T2150CDS  
OPERATING SYSTEM: Windows95

SOFTWARE: Word Perfect 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/064,693A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: W. Gary Goodson  
REGISTRATION NUMBER: 22,387  
REFERENCE/DOCKET NUMBER: LIT-PI-296  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (208)526-9469  
TELEFAX: (208)526-8339  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6641 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-064-693A-25

Query Match 72.8%; Score 498; DB 3; Length 6641;  
Best Local Similarity 100.0%; Pred. No. 2.2e-164;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 187 ATGTACACCCCGGTTAAGCAACTGCGTGGCGCTATTGTTAAAGAGATCCCTGCC 246  
Db 321 ATGTACACCCCGGTTAAGCAACTGCGTGGCGCTATTGTTAAAGAGATCCCTGCC 380  
QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 306  
Db 381 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 440  
QY 307 CAGCAGGGGAAAAACGGTAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGAA 366  
Db 441 CAGCAGGGGAAAAACGGTAAGCGTACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGAA 500  
QY 367 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAAATTTTG 426  
Db 501 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAAATTTTG 560  
QY 427 TTATGTGCGGATGGTGAACCGTGGCTTGCCTGTTACCGTCTGTTTACCGTCTGTTTACCG 486  
Db 561 TTATGTGCGGATGGTGAACCGTGGCTTGCCTGTTACCGTCTGTTTACCGTCTGTTTACCG 620  
QY 487 AGCGGCGCGGAGTGGCGTTTACAAAATTTGGGTAAACCGCGTTAGGAGCGTATCTGTTTC 546  
Db 621 AGCGGCGCGGAGTGGCGTTTACAAAATTTGGGTAAACCGCGTTAGGAGCGTATCTGTTTC 680  
QY 547 ACATCATGACATTAAACCGGAGCTTTATTTAGATAGGCGGCTGATCCGGGCTGTGGGG 606  
Db 681 ACATCATGACATTAAACCGGAGCTTTATTTAGATAGGCGGCTGATCCGGGCTGTGGGG 740  
QY 607 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTACCG 666  
Db 741 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTACCG 800  
QY 667 GCGTCACCGTTGTACTAA 684  
Db 801 GCGTCACCGTTGTACTAA 818

RESULT 5  
US-09-896-866B-12  
Sequence 12, Application US/09896866B  
Patent No. 6642036  
GENERAL INFORMATION:  
APPLICANT: Flint, Dennis  
APPLICANT: Meyer, Knut  
APPLICANT: Viitanen, Paul  
TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate



```
; TITLE OF INVENTION: Benzoin Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-896-866B-12

Query Match      72.4%; Score 495; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.1e-164;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 246
Db 1 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 60

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 120

QY 307 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTCTGAGCAGAAATGA 366
Db 121 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTCTGAGCAGAAATGA 180

QY 367 ATCCCGGAGAACTGCCGCTGCTCCGGAAGAGTCTCGTTACTGGTACGTGAAATTTG 426
Db 181 ATCCCGGAGAACTGCCGCTGCTCCGGAAGAGTCTCGTTACTGGTACGTGAAATTTG 240

QY 427 TTATGTCCGATGTTGAACCGTGGCTTACAAAATTTGGTAAACCGTATCTGTTTC 486
Db 241 TTATGTCCGATGTTGAACCGTGGCTTACAAAATTTGGTAAACCGTATCTGTTTC 300

QY 487 AGCGGGCGGAGCTGGCGTTACAAAATTTGGTAAACCGTATCTGTTTC 546
Db 301 AGCGGGCGGAGCTGGCGTTACAAAATTTGGTAAACCGTATCTGTTTC 360

QY 547 ACATCATCGACATTAACCGGACTTTATTTAGATAGGCGCGTATAGGAGCTATCTGTT 606
Db 361 ACATCATCGACATTAACCGGACTTTATTTAGATAGGCGCGTATAGGAGCTATCTGTT 420

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 666
Db 421 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 480

QY 667 GCGTCACCGTTGTAC 681
Db 481 GCGTCACCGTTGTAC 495
```

```
RESULT 6
US-09-855-341-3
; Sequence 3, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUF
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
```

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: E. coli
US-09-855-341-3

Query Match      72.4%; Score 495; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.1e-164;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 246
Db 1 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 60

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 120

QY 307 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTCTGAGCAGAAATGA 366
Db 121 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCAAGGGTTTGTCTGAGCAGAAATGA 180

QY 367 ATCCCGGAGAACTGCCGCTGCTCCGGAAGAGTCTCGTTACTGGTACGTGAAATTTG 426
Db 181 ATCCCGGAGAACTGCCGCTGCTCCGGAAGAGTCTCGTTACTGGTACGTGAAATTTG 240

QY 427 TTATGTCCGATGTTGAACCGTGGCTTACAAAATTTGGTAAACCGTATCTGTTTC 486
Db 241 TTATGTCCGATGTTGAACCGTGGCTTACAAAATTTGGTAAACCGTATCTGTTTC 300

QY 487 AGCGGGCGGAGCTGGCGTTACAAAATTTGGTAAACCGTATCTGTTTC 546
Db 301 AGCGGGCGGAGCTGGCGTTACAAAATTTGGTAAACCGTATCTGTTTC 360

QY 547 ACATCATCGACATTAACCGGACTTTATTTAGATAGGCGCGTATAGGAGCTATCTGTT 606
Db 361 ACATCATCGACATTAACCGGACTTTATTTAGATAGGCGCGTATAGGAGCTATCTGTT 420

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 666
Db 421 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTACCG 480

QY 667 GCGTCACCGTTGTAC 681
Db 481 GCGTCACCGTTGTAC 495

RESULT 7
US-09-489-039A-1234
; Sequence 1234, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1234
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1234

Query Match      41.2%; Score 282; DB 3; Length 645;
Best Local Similarity 72.9%; Pred. No. 1e-88;
Matches 363; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 246
Db 148 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCGCTATTGTAAGAGAGATCCCTGCC 207
```



247	Qy	CTGGATCCGCAACTGCTCGACTGGCTGTTCTGAGAGTTCATGACAAACGTTTTGAA	306
248	Qy		
208	Db	CTCGCGCGCGCTAAGTGAATGCTGGCTGCTGAGGACTCAATGACCCAGCGCTTTGAA	267
209	Db		
307	Qy	CAGCAGGGAACCGTAAGCGTCACGATATCCGCAAGGTTTTGTGCAGCAGATGAA	366
308	Qy		
268	Db	CAACAGGAAGCAGGTCAACGTGACCTGTTAACGAGGGTATATCGGCGGTGACGGC	327
269	Db		
367	Qy	ATCCCCAAGAACTCCGCTCTCCGCAAGAGTCTCGTTACTGGTTACGTGAAATTTTG	426
368	Qy		
328	Db	CTGACTGACAAAGCGCGCTGTCGCCGACGAGCGCGCTACTGGCTGCGGAGATTATA	387
329	Db		
427	Qy	TTATGTGCGGATGTGAACGTGGCTTGCGGCTCGTACCGTCCTGTGTCAACGTTA	486
428	Qy		
388	Db	CTCAATGCCGATGGCGAGCCCTGGCTGCGCGGGCGACGGTGGCGCGAGTCGACGCTG	447
389	Db		
487	Qy	AGCGGCGGAGCTGCGCTTACAAAAATGGGTAAAAACGCCGTTAGGACGCTATCTGTTC	546
488	Qy		
448	Db	TGTGTTCCGAGCTCGGCTACAGCAGCTCGGCAAACTCGGCTGGSCGATACCTGTTT	507
449	Db		
547	Qy	ACATCATCGACATTAACCGGACTTTATTGAGATAGGCGCGTATGCCGGGCTGTGGGG	606
548	Qy		
508	Db	ACGTCGTGACATTAACCCGCGATTTATTGAAATTGGTCGCGATGACGCGCTGTGGGG	567
509	Db		
607	Qy	CGACGTTTCCGCTCCGATTAAGCGGTAAACCGCTGTTGTAAACAGAACTGTTTTTACCG	666
608	Qy		
568	Db	CGTGGTCCCGCTCGGCTGAGCGGCAACCCCTGCTGTGACCGACTTTTTTTGGCT	627
569	Db		
667	Qy	GGGTCAACGGTTGACTAA	684
668	Qy		
628	Db	GGGTGGCGTTGACTAA	645
629	Db		

```

RESULT 8
US-09-839-477-3
; Sequence 3, Application US/09839477
; Patent No. 6723895
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; TITLE OF INVENTION: COA-CARBOXYLASE
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (204)
US-09-839-477-3

```

```

Db      121  CAAAACCTTGACATCACTCCCATTCGCCAGAACGGCGGAAGAGTGCATATGCATGCAGGTG 180

Qy      181  TGGC 184
          ||||
Db      181  TGGC 184

RESULT 9
US-09-495-797-41
; Sequence 41, Application US/09495797
; Patent No. 6369296
; GENERAL INFORMATION:
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Newburn 43,047
; CURRENT APPLICATION NUMBER: US/09/495,797
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence
US-09-495-797-41

```

```

1  RESULT 10
2  US-08-152-483B-8
3  Sequence 8, Application US/08152483B
4  Patent No. 5529909
5  GENERAL INFORMATION:
6  APPLICANT: Della-Gioppa, Guy
7  APPLICANT: Kumagai, Monto
8  TITLE OF INVENTION: TYROSINASE-ACTIVATOR
9  TITLE OF INVENTION: PROTEIN FUSION ENZYME
10 NUMBER OF SEQUENCES: 9
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Pennie & Edmonds
13 STREET: 2730 Sand Hill Road
14 CITY: Menlo Park
15 STATE: California
16 COUNTRY: U.S.A.
17 ZIP: 94025
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Patent in Release #1.0,
23 SOFTWARE: Version #1.25

```



```
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.24
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/096,043
;; FILING DATE: 22-JUL-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/785,568
;; FILING DATE: 30-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Galloway, No. 5530189val B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 3128567180
;; TELEFAX: 3128564972
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 177 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-096-043-12
;;
Query Match 19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTACACAGCT 60
DB 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTCTCAAGCT 60
QY 61 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
DB 61 AACATGGTGGCGCTTTCACCTGAGCTGAGCTGCTTTCCTCTGTTCACAAAGAAG 120
QY 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
DB 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
;;
RESULT 13
US-08-093-577-8
; Sequence 8, Application US/08093577
; Patent No. 5545816
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Varger, James G
; APPLICANT: Yen, Hwei-Che B
; TITLE OF INVENTION: Phytoene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,577
; FILING DATE: 19-JUL-1993
; CLASSIFICATION: 435
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.24
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/096,043
;; FILING DATE: 22-JUL-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/785,568
;; FILING DATE: 30-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Galloway, No. 5530189val B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 3128567180
;; TELEFAX: 3128564972
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 177 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-096-043-12
;;
Query Match 19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTACACAGCT 60
DB 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTCTCAAGCT 60
QY 61 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
DB 61 AACATGGTGGCGCTTTCACCTGAGCTGAGCTGCTTTCCTCTGTTCACAAAGAAG 120
QY 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
DB 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
;;
RESULT 14
US-08-331-004A-6
; Sequence 6, Application US/08331004A
; Patent No. 5618988
; GENERAL INFORMATION:
; APPLICANT: Hauptmann, Randal
; APPLICANT: Eschenfeldt, William H
; APPLICANT: English, Jami
; APPLICANT: Brinkhaus, Friedhelm L
; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
; TITLE OF INVENTION: in Storage Organs of Genetically
; TITLE OF INVENTION: Engineered Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corporation, Law Dept
; STREET: 55 Shuman Boulevard, Suite 600
; CITY: Naperville
; STATE: IL
; COUNTRY: USA
; ZIP: 60563-8437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,004A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5618988val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7087172447
; TELEFAX: 7087172430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/785,569
;; FILING DATE: 30-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Galloway, No. 5545816val B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 3128567180
;; TELEFAX: 3128564972
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 177 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-093-577-8
;;
Query Match 19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTACACAGCT 60
DB 1 ATGGCTTCCTCTCTATTTCTTTCAGCAGCTGTGGCCACAGCAGCAATGTTCTCAAGCT 60
QY 61 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
DB 61 AACATGGTGGCGCTTTCACCTGAGCTGAGCTGCTTTCCTCTGTTCACAAAGAAG 120
QY 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
DB 121 CAAACCTTGACATCTTCCATTTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
;;
RESULT 14
US-08-331-004A-6
; Sequence 6, Application US/08331004A
; Patent No. 5618988
; GENERAL INFORMATION:
; APPLICANT: Hauptmann, Randal
; APPLICANT: Eschenfeldt, William H
; APPLICANT: English, Jami
; APPLICANT: Brinkhaus, Friedhelm L
; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
; TITLE OF INVENTION: in Storage Organs of Genetically
; TITLE OF INVENTION: Engineered Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corporation, Law Dept
; STREET: 55 Shuman Boulevard, Suite 600
; CITY: Naperville
; STATE: IL
; COUNTRY: USA
; ZIP: 60563-8437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,004A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5618988val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7087172447
; TELEFAX: 7087172430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-331-004A-6

Query Match      19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTCTCAAGCT 60

QY 61 AGCATGGTTCACCTTTTCTCAATCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAG 120
Db 61 AACATGGTGGCGCTTTTCTCAATCTTTCAGCAGCTGTTGCCACACGAGCAATGTTCTCAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTCCTAGCATGTTGGAGAGTAGCTGCATGCAG 177
Db 121 CAAAACCTTGACATCACTTCCATTCCTAGCATGTTGGAGAGTAGCTGCATGCAG 177

RESULT 15
US-08-096-623A-12
; Sequence 12, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Welsh & Katz, Ltd.
; STREET: 120 S. Riverside Plaza, 22nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,623A
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,061
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,921
; FILING DATE: 28-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/562,674
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/525,551
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,613
; FILING DATE: 02-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: AMO-006.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 655-1500
; TELEFAX: (312) 655-1501
```

```
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; MAP POSITION: 1 to 177
; UNITS: bp
US-08-096-623A-12

Query Match      19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTCTCAAGCT 60

QY 61 AGCATGGTTCACCTTTTCTCAATCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAG 120
Db 61 AACATGGTGGCGCTTTTCTCAATCTTTCAGCAGCTGTTGCCACACGAGCAATGTTCTCAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTCCTAGCATGTTGGAGAGTAGCTGCATGCAG 177
Db 121 CAAAACCTTGACATCACTTCCATTCCTAGCATGTTGGAGAGTAGCTGCATGCAG 177
```

Search completed: July 11, 2006, 19:21:21  
Job time : 192 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 19:10:06 ; Search time 4407 Seconds  
(without alignments)  
8679.107 Million cell updates/sec

Title: US-10-718-311-7  
Perfect score: 684  
Sequence: 1 aggtcttcctgtctatttc.....cggcgtcaccgtgtactaa 684

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est3:\*
- 3: gb\_est4:\*
- 4: gb\_est5:\*
- 5: gb\_est6:\*
- 6: gb\_hic:\*
- 7: gb\_est2:\*
- 8: gb\_est7:\*
- 9: gb\_est8:\*
- 10: gb\_est9:\*
- 11: gb\_ges1:\*
- 12: gb\_ges2:\*
- 13: gb\_ges3:\*
- 14: gb\_ges4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	363	53.1	832	13	CL666350 PRI0152b
2	295.4	43.2	828	13	CL687382 PRI0146c
3	184	26.9	205	7	BE462847 EST325226
4	184	26.9	235	7	AW039464 EST281745
5	184	26.9	237	7	AW039041 EST281014
6	184	26.9	238	7	AW041024 EST283888
7	184	26.9	250	7	AW442951 EST307881
8	184	26.9	260	7	AW037699 EST279328
9	184	26.9	260	7	AW094122 EST287302
10	184	26.9	267	1	AW173928 EST255028
11	184	26.9	277	1	AW173502 EST254602
12	184	26.9	279	2	BI931188 EST551077
13	184	26.9	279	7	BE462486 EST324752
14	184	26.9	283	7	AW094091 EST287271
15	184	26.9	286	7	BF051382 EST436557
16	184	26.9	288	7	AW037811 EST279440
17	184	26.9	294	7	AW040243 EST282749
18	184	26.9	329	7	AW038470 EST280153
19	184	26.9	330	7	AW092013 EST285193

20	184	26.9	351	2	BG129648	BG129648	EST475294
21	184	26.9	356	2	BI927635	BI927635	EST547524
22	184	26.9	367	7	AW039845	AW039845	EST282318
23	184	26.9	370	1	AI773548	AI773548	EST254648
24	184	26.9	372	7	BE463004	BE463004	EST325266
25	184	26.9	373	7	AW093805	AW093805	EST286985
26	184	26.9	375	2	BG123527	BG123527	EST469173
27	184	26.9	377	7	AW038793	AW038793	EST280749
28	184	26.9	382	1	AI777051	AI777051	EST252018
29	184	26.9	385	7	AW092154	AW092154	EST285250
30	184	26.9	393	7	AW040522	AW040522	EST283482
31	184	26.9	394	7	BG128649	BG128649	EST474295
32	184	26.9	395	7	AW217063	AW217063	EST295777
33	184	26.9	401	7	AW093581	AW093581	EST286761
34	184	26.9	404	2	BM535307	BM535307	EST588329
35	184	26.9	407	2	BI929701	BI929701	EST549590
36	184	26.9	408	7	AW038880	AW038880	EST280836
37	184	26.9	412	2	BI928909	BI928909	EST548798
38	184	26.9	416	7	AW094716	AW094716	EST287896
39	184	26.9	417	7	AW933080	AW933080	EST358923
40	184	26.9	420	7	AW041064	AW041064	EST283928
41	184	26.9	420	7	AW443412	AW443412	EST308342
42	184	26.9	421	2	BG133589	BG133589	EST466577
43	184	26.9	421	2	BG642684	BG642684	EST510878
44	184	26.9	422	1	AI775340	AI775340	EST256440
45	184	26.9	423	7	AW041189	AW041189	EST284053

## ALIGNMENTS

## RESULT 1

CL666350/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CL666350 832 bp DNA linear GSS 09-JUL-2004  
PRI0152b A03 - PRI0152b.B21 (832) Mixed stage fosmid library of P.  
pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.

CL666350.1 GI:50159025

GSS.

Pristionchus pacificus

Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

1 (bases 1 to 832)

Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.

AppaDB: an AcedB database for the nematode satellite organism

Pristionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

14681447

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: raif.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

Location/Qualifiers

1..832

/organism="Pristionchus pacificus"

/mol\_type="genomic DNA"

/strain="California"

/db\_xref="taxon:54126"

/clone\_lib="Mixed stage fosmid library of P. pacificus

var. California"

/note="Vector: pEpifos-5 Fosmid vector"

## ORIGIN

Query Match 53.1%; Score 363; DB 13; Length 832;

Best Local Similarity 100.0%; Pred. No. 1.8e-104;  
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 GTAAGCGTGACGATGATCCGGAAGGGTTTGTGAGAGAGATGAATCCCGAAGAACTG 381  
|||||  
Db 832 GTAAGCGTGACGATGATCCGGAAGGGTTTGTGAGAGAGATGAATCCCGAAGAACTG 773

QY 382 CGCGTCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAAATTTTGTATGTGCCGATGGT 441  
|||||  
Db 772 CGCGTCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAAATTTTGTATGTGCCGATGGT 713

QY 442 GAACCGTGTGCGTTCGCGGTGATACCGTCTGTTCTGTTCAACGTTTAAGCGGCGCGAGCTG 501  
|||||  
Db 712 GAACCGTGTGCGTTCGCGGTGATACCGTCTGTTCTGTTCAACGTTTAAGCGGCGCGAGCTG 653

QY 502 GCGTTACAAAATTGGGTAAACCGCGTTAGACGCTATCTGTTACATCATCGACATTA 561  
|||||  
Db 652 GCGTTACAAAATTGGGTAAACCGCGTTAGACGCTATCTGTTACATCATCGACATTA 593

QY 562 ACCCGGAGCTTTATTGAGATAGCGGTGATGCGGCGTGTGGGCGGACGTTCCCGCGCTG 621  
|||||  
Db 532 ACCCGGAGCTTTATTGAGATAGCGGTGATGCGGCGTGTGGGCGGACGTTCCCGCGCTG 533

QY 622 CGATTAAAGCGGTAAACCGGTGTTGCTAAACAGAACTGTTTTTACCAGCGTCAACGTTGTAC 681  
|||||  
Db 532 CGATTAAAGCGGTAAACCGGTGTTGCTAAACAGAACTGTTTTTACCAGCGTCAACGTTGTAC 473

QY 682 TAA 684  
|||  
Db 472 TAA 470

RESULT 2  
CL687382  
LOCUS  
DEFINITION  
PR10146c\_H03.2 - PR10146c.BR (828) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Pristionchus pacificus  
Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT

1 (bases 1 to 828)  
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
AppADB: an AcedB database for the nematode satellite organism  
Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
1468147  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.  
Seq primer: 77  
Class: fosmid ends.

FEATURES  
source  
Location/Qualifiers  
1..828  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus var. California"  
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 43.2%; Score 295.4; DB 13; Length 828;  
Best Local Similarity 99.7%; Pred. No. 8.7e-83;  
Matches 296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 187 ATGTACACACCCCGGTAAACGCAACTGCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 246  
|||||  
Db 532 ATGTACACACCCCGGTAAACGCAACTGCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 591

QY 247 CTGGATCCGCAACTGCTGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 306  
|||||  
Db 592 CTGGATCCGCAACTGCTGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 651

QY 307 CAGCAGGAAAAACGGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAGAAATGA 366  
|||||  
Db 652 CAGCAGGAAAAACGGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAGAAATGA 711

QY 367 ATCCCGGAAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAAATTTTG 426  
|||||  
Db 712 ATCCCGGAAGAACTGCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAAATTTTG 771

QY 427 TTATGTGCCGATGTGAACCGTGGCTTCCCGTCTGACCGTTCCTGTGTCAACG 483  
|||||  
Db 772 TTATGTGCCGATGTGAACCGTGGCTTCCCGTCTGACCGTTCCTGTGTCAACG 828

RESULT 3  
BE462847  
LOCUS  
DEFINITION  
EST325226 tomato flower buds 0-3 mm, Cornell University  
Lycopersicon esculentum cDNA clone cfoA16011, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Lycopersicon esculentum (Solanum lycopersicum)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 205)  
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang,P., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
Generation of ESTs from tomato flower tissue, 0-3 mm buds  
Unpublished (1999)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
Location/Qualifiers  
1..205  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cfoA16011"  
/tissue\_type="flower"  
/dev\_stage="0-3mm buds"  
/clone\_lib="tomato flower buds 0-3 mm, Cornell University"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tankley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN

Query Match 26.9%; Score 184; DB 7; Length 205;  
Best Local Similarity 100.0%; Pred. No. 2.9e-47;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTGTGTCATTCTTCAGCAGCTGTTCACGACGCAATGTTACAAAGCT 60

```
|||||
8 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 67
|||||
61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 120
|||||
68 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 127
|||||
121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
|||||
128 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 187
|||||
181 TGGC 184
188 TGGC 191

RESULT 4
AW039464
LOCUS
DEFINITION
EST281745 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET10112, mRNA sequence.
ACCESSION
AW039464
VERSION
AW039464.1 GI:5898218
KEYWORDS
EST.
SOURCE
Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 235)
REFERENCE
1 (bases 1 to 235)
AUTHORS
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE
Generation of ESTs from tomato leaf tissue
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
Location/Qualifiers
source
1..235
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET10112"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="Xl1-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 235;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 60
14 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 73
|||||
61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 120
|||||
74 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 133
|||||
```

```
121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
|||||
134 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 193
|||||
181 TGGC 184
194 TGGC 197

RESULT 5
AW039041
LOCUS
DEFINITION
EST281014 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET10C2, mRNA sequence.
ACCESSION
AW039041
VERSION
AW039041.1 GI:5897795
KEYWORDS
EST.
SOURCE
Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 237)
REFERENCE
1 (bases 1 to 237)
AUTHORS
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE
Generation of ESTs from tomato leaf tissue
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
Location/Qualifiers
source
1..237
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET10C2"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="Xl1-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 237;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 60
9 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 68
|||||
61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 120
|||||
69 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAAGAAG 128
|||||
121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
|||||
129 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 188
|||||
181 TGGC 184
189 TGGC 192
```

```

RESULT 6
AW041024
LOCUS
DEFINITION EST283888 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET615, mRNA sequence.
ACCESSION AW041024
VERSION AW041024.1 GI:5899778
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 238)
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Rönnberg,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tankesley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato leaf tissue
Contact: CUGI
Unpublished (1999)
FEATURES
Location/Qualifiers
1..238
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone="cLET615"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLET - inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisocotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 238;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
|||||
DB 19 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 78
|||||
QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
|||||
DB 79 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 138
|||||
QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGAGGTG 180
|||||
DB 139 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGAGGTG 198
|||||
QY 181 TGGC 184
DB 199 TGGC 202

RESULT 7
AW442951
LOCUS
DEFINITION EST307881 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET42H11 5', mRNA sequence.
ACCESSION AW442951
VERSION AW442951.1 GI:6985133
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 250)
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Hansen,T.S., Rönnberg,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Tankesley,S.D.,
Giovannoni,J. and Martin,G.B.
Generation of ESTs from tomato callus (mixed elicitor)
Unpublished (1999)
CONTACT: CUGI
FEATURES
Location/Qualifiers
1..250
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone="cLET42H11"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLET - inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisocotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 250;
Best Local Similarity 100.0%; Pred. No. 3.1e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
|||||
DB 6 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 65
|||||
QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
|||||
DB 66 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 125
|||||
QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGAGGTG 180
|||||
DB 126 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGAGGTG 185
|||||
QY 181 TGGC 184
DB 186 TGGC 189

RESULT 8
AW037699
LOCUS
DEFINITION EST279328 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET4M17, mRNA sequence.
ACCESSION AW037699
VERSION AW037699.1 GI:5896453
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

```



asteroids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 260)  
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,  
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,  
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and  
 Giovannoni, J.

Generation of ESTs from tomato leaf tissue  
 Unpublished (1999)

CONTACT: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES  
 source  
 1..260  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET4M17"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

ORIGIN  
 Query Match 26.9%; Score 184; DB 7; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-47;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGCAATGTTACCAAGCT 60  
 Db 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGCAATGTTACCAAGCT 71

QY 61 AGCATGGTGGCACCCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120  
 Db 72 AGCATGGTGGCACCCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 131

QY 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180  
 Db 132 CAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 191

QY 181 TGGC 184  
 Db 192 TGGC 195

RESULT 9  
 AW094122 260 bp mRNA linear EST 18-MAY-2001  
 LOCUS  
 DEFINITION  
 EST287302 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 clone cLET27G20, mRNA sequence.

ACCESSION  
 AW094122  
 VERSION  
 AW094122.1 GI:6059717

KEYWORDS  
 EST.

SOURCE  
 Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
 1 (bases 1 to 260)  
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,  
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,  
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and  
 Giovannoni, J.

Generation of ESTs from tomato leaf tissue

TITLE  
 Unpublished (1999)

Unpublished (1999)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES  
 Location/Qualifiers  
 1..260  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET27G20"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

ORIGIN  
 Query Match 26.9%; Score 184; DB 7; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-47;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGCAATGTTACCAAGCT 60  
 Db 6 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGCAATGTTACCAAGCT 65

QY 61 AGCATGGTGGCACCCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120  
 Db 66 AGCATGGTGGCACCCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 125

QY 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180  
 Db 126 CAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 185

QY 181 TGGC 184  
 Db 186 TGGC 189

RESULT 10  
 AW094122 260 bp mRNA linear EST 18-MAY-2001  
 LOCUS  
 DEFINITION  
 EST255028 tomato resistant, Cornell Lycopersicon esculentum cDNA  
 clone cLER8L7, mRNA sequence.

ACCESSION  
 AW094122  
 VERSION  
 AW094122.1 GI:5271969

KEYWORDS  
 EST.

SOURCE  
 Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
 1 (bases 1 to 267)  
 D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,  
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,  
 Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,  
 Giovannoni, J.J. and Martin, G.B.

Generation of ESTs from Pseudomonas resistant tomato

TITLE  
 Unpublished (1999)

JOURNAL  
 COMMENT  
 Contact: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

```

FEATURES
source
Location/Qualifiers
1..267
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S)::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER817"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Query Match 26.9%; Score 184; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.2e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
Db 23 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 82

QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
Db 83 AGCATGGTTGCACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 142

QY 121 CAAACCTTGACATCATTCCATTCCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db 143 CAAACCTTGACATCATTCCATTCCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 202

QY 181 TGGC 184
Db 203 TGGC 206

RESULT 11
A1773502
LOCUS A1773502 277 bp mRNA linear EST 18-MAY-2001
DEFINITION EST254602 tomato resistant, Cornell Lycopersicon esculentum cDNA
clone CLER7A10, mRNA sequence.
ACCESSION A1773502
VERSION A1773502.1 GI:5271543
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 277)
REFERENCE 1
AUTHORS D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
Bowman, C.L., Niemman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J. and Martin, G.B.
TITLE Generation of ESTs from Pseudomonas resistant tomato
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..277
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S)::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER7A10"
/tissue_type="leaf"
/dev_stage="4-week old"

FEATURES
source
Location/Qualifiers
1..279
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0C19C6"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Query Match 26.9%; Score 184; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.2e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
Db 5 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 64

QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
Db 65 AGCATGGTTGCACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 124

QY 121 CAAACCTTGACATCATTCCATTCCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db 125 CAAACCTTGACATCATTCCATTCCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 184

QY 181 TGGC 184
Db 185 TGGC 188

RESULT 12
BI931188
LOCUS BI931188 279 bp mRNA linear EST 18-OCT-2001
DEFINITION EST551077 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cT0C19C6 5' end, mRNA sequence.
ACCESSION BI931188
VERSION BI931188.1 GI:16245660
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 279)
REFERENCE 1
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
Utterback, T., Van Aken, S., Ronning, C.M., Niemman, W., Fraser, C.M.,
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
FEATURES
source
Location/Qualifiers
1..279
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0C19C6"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

## ORIGIN

Query Match 26.9%; Score 184; DB 2; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 60  
 DB 7 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 66  
 QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120  
 DB 67 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 126  
 QY 121 CAAAACCTTGACATCATTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTTG 180  
 DB 127 CAAAACCTTGACATCATTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTTG 186  
 QY 181 TGGC 184  
 DB 187 TGGC 190

## RESULT 13

BE462486  
 LOCUS  
 DEFINITION EST324752 tomato flower buds 0-3 mm, Cornell University  
 Lycopersicon esculentum cDNA clone cTOAL301, mRNA sequence.

ACCESSION BE462486  
 VERSION BE462486.1  
 KEYWORDS GI:9508255

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

1 (bases 1 to 279)  
 van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E.,  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and  
 Tanksley, S.D.

## TITLE

Generation of ESTs from tomato flower tissue, 0-3 mm buds

## JOURNAL

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

Location/Qualifiers

1..279  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOAL301"  
 /tissue\_type="flower"  
 /dev\_stage="0-3mm buds"  
 /clone\_lib="tomato flower buds 0-3 mm, Cornell University"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."

## ORIGIN

Query Match 26.9%; Score 184; DB 7; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 60  
 DB 7 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 66

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120  
 DB 67 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 126  
 QY 121 CAAAACCTTGACATCATTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTTG 180  
 DB 127 CAAAACCTTGACATCATTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTTG 186  
 QY 181 TGGC 184  
 DB 187 TGGC 190

## RESULT 14

AW094091  
 LOCUS  
 DEFINITION EST287271 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA  
 clone cLET27A10, mRNA sequence.

ACCESSION AW094091  
 VERSION AW094091.1  
 KEYWORDS GI:6059686

## SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

1 (bases 1 to 283)

## AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,  
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,  
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and  
 Giovannoni, J.J.

Generation of ESTs from tomato leaf tissue

## TITLE

Unpublished (1999)

## JOURNAL

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

Location/Qualifiers

1..283  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="Rio Grande PtoR"  
 /db\_xref="taxon:4081"  
 /clone="cLET27A10"  
 /tissue\_type="leaf"  
 /dev\_stage="4-6 week old plants"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato mixed elicitor, BTI"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLET - Inoculated with a variety of disease response  
 elicitors. Plants exposed to 2,6 dichloroisonicotinic  
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,  
 okadaic acid, or systemin prior to tissue harvest. EcoRI  
 site was destroyed during cloning."

## ORIGIN

Query Match 26.9%; Score 184; DB 7; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 60  
 DB 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCACAGCAGCAATGTTACACAAGCT 71  
 QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120  
 DB 72 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 131  
 QY 121 CAAAACCTTGACATCATTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCAGGTTG 180

Db 132 CAAACCTTGACATCACTTCCATTGCTAGCAATGTTGAAGAGTTAGTCGATGCAGGTG 191  
QY 181 TGGC 184  
||||  
Db 192 TGGC 195

RESULT 15  
BF051382  
LOCUS BF051382 286 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST436557 tomato developing/immature green fruit Lycopersicon  
esculentum cDNA clone cLEM22P19 5' sequence, mRNA sequence.  
ACCESSION BF051382  
VERSION BF051382.1 GI:10805278  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 286)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tankley,S.D.  
REFERENCE Generation of ESTs from tomato fruit tissue, immature green  
AUTHORS Unpublished (2000)  
TITLE Contact: CUGI  
JOURNAL Clemson University Genomics Institute  
COMMENT Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
Location/Qualifiers

FEATURES  
source  
1..286  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEM22P19"  
/tissue\_type="fruit"  
/dev\_stage="immature green (5-35 days post-anthesis)"  
/lab\_host="SOLR"  
/clone\_lib="tomato developing/immature green fruit"  
/note="Vector: pBluescriptSKmCUadapt; Site 1: EcoRI;  
Site 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and  
harvested at 7 day intervals through 35 dpa. Equal masses  
of tissue from each stage were combined (including seeds  
and locules) prior to mRNA isolation."

ORIGIN  
Query Match 26.9%; Score 184; DB 7; Length 286;  
Best Local Similarity 100.0%; Pred. No. 3.3e-47;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCCTGTGTCATTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60  
Db |||||||  
10 ATGGCTTCCTGTGTCATTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 69  
QY 61 AGCATGGTTGACCTTCACCTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 120  
Db |||||||  
70 AGCATGGTTGACCTTCACCTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 129  
QY 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGTTGAAGAGTTAGTCGATGCAGGTG 180  
Db |||||||  
130 CAAACCTTGACATCACTTCCATTGCTAGCAATGTTGAAGAGTTAGTCGATGCAGGTG 189  
QY 181 TGGC 184  
||||  
Db 190 TGGC 193

Search completed: July 11, 2006, 21:36:48  
Job time : 4410 secs

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 18:50:32 ; Search time 588 Seconds  
 (without alignments)  
 8110.575 Million cell updates/sec

Title: US-10-718-311-7  
 Perfect score: 684  
 Sequence: 1 atggctctctctgtcatttc.....cggcgtcaccgtgtactaa 684

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N Geneseq 8.\*  
 1: Geneseqn1980s.\*  
 2: Geneseqn1990s.\*  
 3: Geneseqn2000s.\*  
 4: Geneseqn2001as.\*  
 5: Geneseqn2001bs.\*  
 6: Geneseqn2002as.\*  
 7: Geneseqn2002bs.\*  
 8: Geneseqn2003as.\*  
 9: Geneseqn2003bs.\*  
 10: Geneseqn2003cs.\*  
 11: Geneseqn2003ds.\*  
 12: Geneseqn2004as.\*  
 13: Geneseqn2004bs.\*  
 14: Geneseqn2005s.\*  
 15: Geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	684	100.0	684	ABA91838	Chloroplast
2	684	100.0	684	ABA91838	Chloroplast
3	684	100.0	684	ABA91838	Chloroplast
4	684	100.0	684	ABA91838	Chloroplast
5	501	73.2	3452	ADV91636	C. paraps
6	499.4	73.0	1971	ADV91643	C. paraps
7	498	72.8	498	ADV91618	Escherich
8	498	72.8	854	AAD06956	ubc gene
9	498	72.8	6641	AAD06957	pME2 comp
10	496.4	72.6	2000	AAT29820	E. coli u
11	495	72.4	495	AAQ92409	E. coli ch
12	495	72.4	495	ABA91837	Escherich
13	495	72.4	495	ABA91837	Escherich
14	495	72.4	495	ABA91837	Escherich
15	472.8	69.1	1207	AAQ90335	DNA encod
16	472.8	69.1	1207	AAQ90335	DNA encod
17	282	41.2	645	ACH95439	Klebsiell
18	145.6	21.3	492	ADC76164	DNA homol

19	144	21.1	581	10	ADC75089	N bentham
20	144	21.1	614	10	ADC76956	DNA homol
21	144	21.1	632	10	ADC76953	DNA homol
22	144	21.1	718	10	ADK58382	Plant DNA
23	144	21.1	736	10	ADK54321	Plant DNA
24	144	21.1	736	10	ADK57660	Plant DNA
25	144	21.1	841	11	ADM44879	Insect re
26	144	21.1	847	11	ADM45447	Insect re
27	144	21.1	924	10	ADC75566	DNA homol
28	144	21.1	958	10	ADC76165	DNA homol
29	142.4	20.8	297	10	ADK59826	Plant DNA
30	142.4	20.8	377	10	ADK56121	Plant DNA
31	142.4	20.8	489	10	ADK56120	Plant DNA
32	142.4	20.8	504	10	ADK59824	Plant DNA
33	142.4	20.8	608	10	ADC76944	DNA homol
34	142.4	20.8	615	10	ADC76949	DNA homol
35	142.4	20.8	684	10	ADC76948	DNA homol
36	142.4	20.8	859	11	ADM45065	Insect re
37	141.4	20.7	619	11	ADM45158	Insect re
38	140.8	20.6	204	6	AAI70688	Tobacco R
39	140.8	20.6	499	6	AAI70688	Tobacco R
40	140.8	20.6	609	10	ADC76965	Nicotiana
41	137.6	20.1	668	10	ADK59825	Plant DNA
42	137.6	20.1	668	11	ADM45687	Insect re
43	133.2	19.5	1442	2	AAQ92327	Chloropla
44	132.6	19.4	171	14	ABE04209	Tomato rb
45	132.6	19.4	171	14	ABE04212	Tomato rb

## ALIGNMENTS

## RESULT 1

ABA91838  
 ID ABA91838 standard; DNA; 684 BP.

XX ABA91838;  
 AC ABA91838;

XX 29-AUG-2003 (revised)  
 DT 15-MAY-2002 (first entry)

XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.

XX Chloroplast transit peptide; tomato; Rubisco; plant;

KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;  
 KW enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

XX Chimeric.

Key Location/Qualifiers

CDS 1..495

FT /\*tag= a

FT /product= "chloroplast-targeted CPL fusion"

FT transit\_peptide 1..186

FT /\*tag= b

FT mat\_peptide 187..492

FT /\*tag= c

FT /note= "E. coli chorismate pyruvate lyase"

XX WO200194607-A2.

PN 13-DEC-2001.

XX 22-MAY-2001; 2001WO-US016661.

XX 02-JUN-2000; 2000US-0209854P.

XX (DUFO) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;



```
Query Match      100.0%; Score 684; DB 6; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.4e-219;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGCTTCTCTGCTCATTTCTTCAGCAGCTGTTGCCACACGCAATGTTACACAAGCT 60
DB      1 ATGGCTTCTCTGCTCATTTCTTCAGCAGCTGTTGCCACACGCAATGTTACACAAGCT 60

QY      61 AGCATGGTTGCACCTTTCACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB      61 AGCATGGTTGCACCTTTCACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

QY      121 CAARACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180
DB      121 CAARACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180

QY      181 TGGCATATGTCAACACCCCGCTTAACGCAACTGCGTGCCTGCTATTGTTAAAGAGATC 240
DB      181 TGGCATATGTCAACACCCCGCTTAACGCAACTGCGTGCCTGCTATTGTTAAAGAGATC 240

QY      241 CTGCGCTGGATCCGCAACTGCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB      241 CTGCGCTGGATCCGCAACTGCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY      301 TTTGAACAGCAGGGAAGAAACGGTAACGATGATCGCGAAGGGTTGTCGAGCAG 360
DB      301 TTTGAACAGCAGGGAAGAAACGGTAACGATGATCGCGAAGGGTTGTCGAGCAG 360

QY      361 AATGAAATCCCGAAGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB      361 AATGAAATCCCGAAGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY      421 ATTTGTTATGTGCGCATGGTGAACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB      421 ATTTGTTATGTGCGCATGGTGAACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY      481 AGCTTAAGCGGCGCGAGCTGGCTTACAAAATTTGGGTAAACCGCGTTAGACGCTAT 540
DB      481 AGCTTAAGCGGCGCGAGCTGGCTTACAAAATTTGGGTAAACCGCGTTAGACGCTAT 540

QY      541 CTGTTTCAATCATGACATTAACCGGAGCTTTATTGAGATAGGCGGTGATGCGGGCTG 600
DB      541 CTGTTTCAATCATGACATTAACCGGAGCTTTATTGAGATAGGCGGTGATGCGGGCTG 600

QY      601 TGGGGCGAGCTTCCCGCTCGGATTAACGGGTAAACCCGCTGTTGCTAAACAGAACTGTTT 660
DB      601 TGGGGCGAGCTTCCCGCTCGGATTAACGGGTAAACCCGCTGTTGCTAAACAGAACTGTTT 660

QY      661 TTACCGGCGTCACTGTTGACTAA 684
DB      661 TTACCGGCGTCACTGTTGACTAA 684
```

## RESULT 3

```
ADC25988
ID   ADC25988 standard; DNA; 684 BP.
XX
AC   ADC25988;
XX
DT   18-DEC-2003 (first entry)
XX
DE   Tomato/Escherichia coli chimeric TP-CPL DNA.
XX
KW   UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;
KW   liquid crystal polymer; LCP; methylparaben; preservative; food;
KW   cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;
KW   chorismate pyruvate lyase; plant.
XX
OS   Chimeric.
OS   Lycopersicon esculentum.
OS   Escherichia coli.
XX
```

```
FH Key Location/Qualifiers
FT CDS 1..684
FT /tag= a
FT /product= "Tomato/Escherichia coli chimeric TP-CPL
FT protein"
XX
XX WO2003066836-A2.
XX
XX 14-AUG-2003.
XX
XX 06-FEB-2003; 2003WO-US005863.
XX
XX 07-FEB-2002; 2002US-0355511P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Meyer K, Van Dyk DE, Viitanen PV;
XX
XX WPI: 2003-767259/72.
XX
XX P-PSDB; ADC25989.
XX
XX New nucleic acid encoding UDP-glucosyltransferase, useful for preparing
XX cells that produce p-hydroxybenzoic acid glucose ester, also the new
XX enzymes.
XX
XX Example 9; SEQ ID NO 41; 161pp; English.
XX
XX The invention relates to a novel isolated nucleic acid that encodes a UDP
XX -glucosyltransferase. The method of the invention may be used to
XX transform microorganisms or green plant cells so that these produce a
XX higher level of high-volume chemicals or materials, such as pHBA (p-
XX hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a
XX monomer for liquid crystal polymers (LCPs) and starting material for
XX methylparaben. Methylparaben is a preservative commonly used in the food
XX and cosmetic industries. The encoded enzymes may be used for in vitro
XX production of these compounds and for identifying similar enzymes by
XX sequence comparison. The current sequence is that of the
XX tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA
XX of the invention.
XX
XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;
XX
XX Query Match      100.0%; Score 684; DB 10; Length 684;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-219;
XX Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGCTTCTCTGCTCATTTCTTCAGCAGCTGTTGCCACACGCAATGTTACACAAGCT 60
DB      1 ATGGCTTCTCTGCTCATTTCTTCAGCAGCTGTTGCCACACGCAATGTTACACAAGCT 60

QY      61 AGCATGGTTGCACCTTTCACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB      61 AGCATGGTTGCACCTTTCACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

QY      121 CAARACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180
DB      121 CAARACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180

QY      181 TGGCATATGTCAACACCCCGCTTAACGCAACTGCGTGCCTGCTATTGTTAAAGAGATC 240
DB      181 TGGCATATGTCAACACCCCGCTTAACGCAACTGCGTGCCTGCTATTGTTAAAGAGATC 240

QY      241 CTGCGCTGGATCCGCAACTGCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB      241 CTGCGCTGGATCCGCAACTGCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY      301 TTTGAACAGCAGGGAAGAAACGGTAACGATGATCGCGAAGGGTTGTCGAGCAG 360
DB      301 TTTGAACAGCAGGGAAGAAACGGTAACGATGATCGCGAAGGGTTGTCGAGCAG 360

QY      361 AATGAAATCCCGAAGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB      361 AATGAAATCCCGAAGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
```







XX	SQ	Sequence 3452 BP; 932 A; 730 C; 856 G; 934 T; 0 U; 0 Other;
XX	ID	ADV91636 standard; DNA; 3452 BP.
XX	AC	ADV91636;
XX	DT	10-MAR-2005 (first entry)
XX	DE	C. parapsilosis pHB 1-H- E. coli CPL- A. thaliana UGT72B1 chimeric DNA.
XX	KW	Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
XX	KW	chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
XX	KW	HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
XX	KW	genetically engineered microorganism; antioxidant; antimicrobial agent;
XX	KW	anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
XX	KW	neoplasm; pHBA 1-hydroxylase; pHBA 1-H; pHBA; para-hydroxybenzoic acid;
XX	KW	p-hydroxybenzoic acid; UDP-glucosyltransferase; UGT72B1; gene; ds.
XX	OS	Arabidopsis thaliana.
XX	OS	Candida parapsilosis.
XX	OS	Escherichia coli.
XX	OS	Chimeric.
PH	Key	Location/Qualifiers
FT	CDS	4..501
FT	/tag= a	/product= "Escherichia coli CPL protein"
FT	541..1980	
FT	/tag= b	
FT	/product= "Candida parapsilosis pHBA 1-H protein"	
FT	/transl_except= (pos: 559..561, aa:Trp)	
FT	2004..3446	
FT	/tag= c	
FT	/product= "Arabidopsis thaliana UGT72B1 protein"	
XX	PN	US2004261147-A1.
XX	PD	23-DEC-2004.
XX	PJ	16-JUN-2003; 2003US-00462162.
PF	PJ	16-JUN-2003; 2003US-00462162.
PR	PA	(MEYE/) MEYER K.
PA	PA	(VIIT/) VIITANEN P V.
PA	PA	(FLIN/) FLINT D.
PI	PI	Meyer K, Viitanen PV, Flint D;
PI	WIPI	2005-057232/06.
DR	R-PSDB	ADV91605, ADV91619, ADV91631.
XX	PT	Producing hydroquinone glucoside in a green plant comprises growing a
PT	PT	green plant having nucleic acid fragments and chorismate pyruvate lyase
PT	PT	expression cassette.
PS	Example 7	; SEQ ID NO 47; 70pp; English.
CC	CC	The present invention relates to methods and materials to produce
CC	CC	hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)
CC	CC	in genetically modified green plants and microorganisms. The method
CC	CC	relies upon transgenic plants or genetically modified microorganisms that
CC	CC	produce increased levels of the initial substrate para-hydroxybenzoic
CC	CC	acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce
CC	CC	arbutin. Plants and microbes can be genetically engineered to produce
CC	CC	high levels of pHBA either by functional expression of the bacterial
CC	CC	protein chorismate pyruvate lyase (CPL) or by the expression of 4-
CC	CC	hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.
CC	CC	Arbutin is useful as an antioxidant, antimicrobial agent, anti-
CC	CC	inflammatory agent and inhibitor of carcinogenesis (melanoma). The
CC	CC	present sequence is Candida parapsilosis pHBA 1-hydroxylase (pHB 1-H) -
CC	CC	Escherichia coli chorismate pyruvate lyase (CPL) and Arabidopsis thaliana
CC	CC	UDP-glucosyltransferase (UGT72B1) chimeric DNA.

XX	SQ	Sequence 3452 BP; 932 A; 730 C; 856 G; 934 T; 0 U; 0 Other;
XX	ID	ADV91636 standard; DNA; 3452 BP.
XX	AC	ADV91636;
XX	DT	10-MAR-2005 (first entry)
XX	DE	C. parapsilosis pHB 1-H- E. coli CPL- A. thaliana UGT72B1 chimeric DNA.
XX	KW	Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
XX	KW	chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
XX	KW	HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
XX	KW	genetically engineered microorganism; antioxidant; antimicrobial agent;
XX	KW	anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
XX	KW	neoplasm; pHBA 1-hydroxylase; pHBA 1-H; pHBA; para-hydroxybenzoic acid;
XX	KW	p-hydroxybenzoic acid; UDP-glucosyltransferase; UGT72B1; gene; ds.
XX	OS	Arabidopsis thaliana.
XX	OS	Candida parapsilosis.
XX	OS	Escherichia coli.
XX	OS	Chimeric.
PH	Key	Location/Qualifiers
FT	CDS	4..501
FT	FT	/tag= a
FT	FT	/product= "Escherichia coli CPL protein"
FT	CDS	541..1980
FT	FT	/tag= b
FT	FT	/product= "Candida parapsilosis pHBA 1-H protein"
FT	FT	/transl_except= (pos: 559..561, aa:Trp)
FT	CDS	2004..3446
FT	FT	/tag= c
FT	FT	/product= "Arabidopsis thaliana UGT72B1 protein"
XX	PX	US2004261147-A1.
XX	PD	23-DEC-2004.
XX	PF	16-JUN-2003; 2003US-00462162.
XX	PR	16-JUN-2003; 2003US-00462162.
PA	(MEYE/) MEYER K.	
PA	(VIIT/) VIITANEN P V.	
PA	(FLIN/) FLINT D.	
PI	Meyer K, Viitanen PV, Flint D;	
XX	WIPI	2005-057232/06.
DR	P-PSDB; ADV91605, ADV91619, ADV91631.	
XX	PT	Producing hydroquinone glucoside in a green plant comprises growing a
PT	green plant having nucleic acid fragments and chorismate pyruvate lyase	
PT	expression cassette.	
XX	Example 7; SEQ ID NO 47; 70pp; English.	
CC	The present invention relates to methods and materials to produce	
CC	hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)	
CC	in genetically modified green plants and microorganisms. The method	
CC	relies upon transgenic plants or genetically modified microorganisms that	
CC	produce increased levels of the initial substrate para-hydroxybenzoic	
CC	acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce	
CC	arbutin. Plants and microbes can be genetically engineered to produce	
CC	high levels of pHBA either by functional expression of the bacterial	
CC	protein chorismate pyruvate lyase (CPL) or by the expression of 4-	
CC	hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.	
CC	Arbutin is useful as an antioxidant, antimicrobial agent, anti-	
CC	inflammatory agent and inhibitor of carcinogenesis (melanoma). The	
CC	present sequence is Candida parapsilosis pHBA 1-hydroxylase (pHB 1-H) -	
CC	Escherichia coli chorismate pyruvate lyase (CPL) and Arabidopsis thaliana	
CC	UDP-glucosyltransferase (UGT72B1) chimeric DNA.	

```
FT /product= "Candida parapsilosis pHBA 1-H protein"  
FT /transl_except= (pos: 540..542, aa:Trp)  
XX  
PN US2004261147-A1.  
XX 23-DEC-2004.  
XX 16-JUN-2003; 2003US-00462162.  
XX 16-JUN-2003; 2003US-00462162.  
XX (MEYE/) MEYER K.  
PA (VIIT/) VIITANEN P V.  
PA (FLIN/) FLINT D.  
XX  
PI Meyer K, Viitanen PV, Flint D;  
XX WPI; 2005-057232/06.  
DR P-PSDB; ADV91605, ADV91619.  
XX  
PT Producing hydroquinone glucoside in a green plant comprises growing a  
PT green plant having nucleic acid fragments and chorismate pyruvate lyase  
PT expression cassette.  
XX  
XX Example 7; SEQ ID NO 54; 70pp; English.  
XX  
CC The present invention relates to methods and materials to produce  
CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
CC in genetically modified green plants and microorganisms. The method  
CC relies upon transgenic plants or genetically modified microorganisms that  
CC produce increased levels of the initial substrate para-hydroxybenzoic  
CC acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce  
CC arbutin. Plants and microbes can be genetically engineered to produce  
CC high levels of pHBA either by functional expression of the bacterial  
CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-  
CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
CC present sequence is Candida parapsilosis pHBA 1-hydroxylase (pHB 1-H) and  
CC Escherichia coli chorismate pyruvate lyase (CPL) chimeric DNA.  
XX  
SQ Sequence 1971 BP; 550 A; 383 C; 484 G; 554 T; 0 U; 0 Other;  
  
Query Match  
Best Local Similarity 73.0%; Score 499.4; DB 14; Length 1971;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 184 CATATGTACACCCCGGTTAAACGGTAAACGCGTGTGCGTATTTGTAAGAGATCCCT 243  
Db 1 CATATGTACACCCCGGTTAAACGGTAAACGCGTGTGCGTATTTGTAAGAGATCCCT 60  
  
QY 244 GCCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCATGACAAAAGCTTTT 303  
Db 61 GCCCTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCATGACAAAAGCTTTT 120  
  
QY 304 GAACAGCAGGAAAAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAAAC 363  
Db 121 GAACAGCAGGAAAAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAAAC 180  
  
QY 364 GAAATCCCGAAGAACTGCGCGTGTGCGGAAAGATCTCGTTACTGTTAGTGAAATTT 423  
Db 181 GAAATCCCGAAGAACTGCGCGTGTGCGGAAAGATCTCGTTACTGTTAGTGAAATTT 240  
  
QY 424 TTGTTATGTCGGGATGTTGAACCGTGGCTGTCGGTGTGCGGTGCTGCTGTTGTCACAG 483  
Db 241 TTGTTATGTCGGGATGTTGAACCGTGGCTGTCGGTGTGCGGTGCTGCTGTTGTCACAG 300  
  
QY 484 TTAAGCGGCGCGAGCTGCGGTTACAAAAATTTGGGTAAACCGGTTAGGACGCTATCTG 543  
Db 301 TTAAGCGGCGCGAGCTGCGGTTACAAAAATTTGGGTAAACCGGTTAGGACGCTATCTG 360  
  
QY 544 TTCACATCATCACATTAAACCGGACCTTTATTGAGATAGGCCGTGATCCCGGCTGTGG 603  
|||||
```

```
Db 361 TTCACATCATCGACATTAAACCGGACCTTTATTGAGATAGGCCGTGATCCCGGCTGTGG 420  
QY 604 GGGGACGCTCCCGCTCGGATTAAACCGGTTAAACCGCTGTTGCTTAACAGAACTGTTTTTA 663  
Db 421 GGGGACGCTCCCGCTCGGATTAAACCGGTTAAACCGCTGTTGCTTAACAGAACTGTTTTTA 480  
QY 664 CCGGCTCACCGTTGTACTAA 684  
Db 481 CCGGCTCACCGTTGTACTAA 501  
  
RESULT 7  
ADV91618  
ID ADV91618 standard; DNA; 498 BP.  
XX  
XX ADV91618;  
AC  
XX 10-MAR-2005 (first entry)  
DT  
XX  
DE Escherichia coli chorismate pyruvate lyase (CPL) DNA.  
XX  
KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;  
KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;  
KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;  
KW genetically engineered microorganism; antioxidant; antimicrobial agent;  
KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;  
KW neoplasm; pHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid; gene;  
KW ds.  
XX  
OS Escherichia coli.  
FH Key Location/Qualifiers  
CDS 1..498  
FT /*tag= a  
FT /product= "Escherichia coli CPL protein"  
FT  
XX US2004261147-A1.  
PD 23-DEC-2004.  
XX  
XX 16-JUN-2003; 2003US-00462162.  
XX  
XX 16-JUN-2003; 2003US-00462162.  
PA (MEYE/) MEYER K.  
PA (VIIT/) VIITANEN P V.  
PA (FLIN/) FLINT D.  
XX  
XX Meyer K, Viitanen PV, Flint D;  
XX WPI; 2005-057232/06.  
DR P-PSDB; ADV91619.  
XX  
XX Producing hydroquinone glucoside in a green plant comprises growing a  
XX green plant having nucleic acid fragments and chorismate pyruvate lyase  
XX expression cassette.  
XX  
XX Example 4; SEQ ID NO 29; 70pp; English.  
XX  
CC The present invention relates to methods and materials to produce  
CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)  
CC in genetically modified green plants and microorganisms. The method  
CC relies upon transgenic plants or genetically modified microorganisms that  
CC produce increased levels of the initial substrate para-hydroxybenzoic  
CC acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce  
CC arbutin. Plants and microbes can be genetically engineered to produce  
CC high levels of pHBA either by functional expression of the bacterial  
CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-  
CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.  
CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-  
CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The  
CC present sequence is Escherichia coli chorismate pyruvate lyase (CPL) DNA.  
CC This sequence is found in the E. coli expression vector construct
```







DT	15-MAY-2002 (first entry)	DT	61	CTGGATCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATCGACAAACGTTTGA	120
XX	Escherichia coli chorismate pyruvate lyase gene coding region.	Db			
DE					
XX	Chorismate pyruvate lyase; CPL; enzyme; p-hydroxybenzoic acid;	QY	307	CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGTTTGTGACGACAGATGAA	366
KW	transgenic plant; ubiC; gene; ds.	Db	121	CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGTTTGTGACGACAGATGAA	180
XX	Escherichia coli.	QY	367	ATCCCGGAAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTTGAAATTTTG	426
XX		Db	181	ATCCCGGAAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTTGAAATTTTG	240
FH	Key	QY	427	TTATGTGCCGATGCTGAACCCGCTGCTGCCGTTCCCGTACCGTCTGTTCTGTTCAACGTTA	486
FT	CDS	Db	241	TTATGTGCCGATGCTGAACCCGCTGCTGCCGTTCCCGTACCGTCTGTTCTGTTCAACGTTA	300
FT	1..495	QY	487	AGCGGGCGGAGCTGGCGTTACAAAATTGGTAAACCGCTTAGGACGCTATCTGTTTC	546
FT	/*tag= a	Db	301	AGCGGGCGGAGCTGGCGTTACAAAATTGGTAAACCGCTTAGGACGCTATCTGTTTC	360
FT	/product= "Chorismate pyruvate lyase"	QY	547	ACATCATCGACATTAACCCGGGACTTTATTGAGATAGCCCGTGTATGCCGGGCTGTGGGG	606
FT	/note= "the CDS does not include a stop codon"	Db	361	ACATCATCGACATTAACCCGGGACTTTATTGAGATAGCCCGTGTATGCCGGGCTGTGGGG	420
XX	WO200194607-A2.	QY	607	CGAGTTTCCCGCTGCGGATTAAAGCGGTTAAACCGCTGTTGCTAAACAGAACTGTTTACCG	666
XX	13-DEC-2001.	Db	421	CGAGTTTCCCGCTGCGGATTAAAGCGGTTAAACCGCTGTTGCTAAACAGAACTGTTTACCG	480
XX	22-MAY-2001; 2001WO-US016661.	QY	667	GCCTCACCGTTGTAC	681
XX	02-JUN-2000; 2000US-0209854P.	Db	481	CGCTCACCGTTGTAC	495
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.				
XX	Meyer K, Van Dyk DE, Viitanen PV;				
XX	WPI; 2002-226795/28.				
DR	P-PSDB; AAM50958.				
XX	Producing para-hydroxy benzoic acid in green plant, comprises expression				
PT	of unique expression cassette containing gene encoding chorismate				
PT	pyruvate lyase operably linked to specific chloroplast targeting				
PT	sequence.				
XX					
PS	Claim 4; Page 55; 60pp; English.				
XX	The present sequence is that of the coding region of the Escherichia coli				
CC	strain W3110 ubiC gene in expression construct pET24a. The gene encodes				
CC	chorismate pyruvate lyase (CPL, see AAM50968). It was obtained by PCR				
CC	amplification of strain W3110 genomic DNA using primers (see ABA91839-40)				
CC	based on the published E. coli ubiC gene. The CPL open reading frame is				
CC	used in an expression cassette designed for the high-level production of				
CC	p-hydroxybenzoic acid (pHBA) in green plants. The expression cassette				
CC	comprises the CPL coding sequence operably linked to a promoter capable				
CC	of driving protein expression in higher plants. The cassette also has a				
CC	sequence encoding a chloroplast transit peptide, its natural cleavage				
CC	site, and a small portion of a transit peptide donor protein fused to the				
CC	N-terminus of CPL. The chloroplast targeting sequence targets the foreign				
CC	protein to the chloroplast and aids in its uptake into the organelle. The				
CC	cleavage site is unique to the transit peptide, and cleavage of the				
CC	chimeric protein encoded by the cassette at this site releases a novel				
CC	polypeptide that has full enzyme activity, comprising the mature CPL				
CC	enzyme and a small portion of transit peptide donor (see AAM50961). A				
CC	plant comprising the CPL expression cassette is claimed, and may be				
CC	soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,				
CC	barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,				
CC	millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric				
CC	component of liquid crystal polymers which have application in the				
CC	automotive, electrical and other industries				
XX					
SQ	Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;				
	Query Match				
	Best Local Similarity 72.4%; Score 495; DB 6; Length 495;				
	Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	187 ATGTCACACCCCGGTTAACGCAACTGCGTGGCTGCTATTGTAAGAGATCCCTGCC	246			
Db	1 ATGTCACACCCCGGTTAACGCAACTGCGTGGCTGCTATTGTAAGAGATCCCTGCC	60			
QY	247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATCGACAAACGTTTGA	306			

Producing aromatic acid conjugates, involves contacting glycosylated

301 AGCGGCCGGAGCTGGCGTTACAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTCTC 360

UDP-glucosyltransferase; PHBA; p-hydroxybenzoic acid ester glucoside;  
 KW KW liquid crystal polymer; LCP; methylparaben; preservative; food;  
 KW KW cosmetic industry; ds; gene; ubiC; CPL; chorismate pyruvate lyase.



QY 547 ACATCATGACATTAAACCGGACCTTATTGAGATAGCCGCTGATGCCGGCTGTGGGG 606  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 ACATCATGACATTAAACCGGACCTTATTGAGATAGCCGCTGATGCCGGCTGTGGGG 420  
QY 607 CGACGTTCCCGCTCGATTAAACCGGTAACCGCTGTTGCTAACAGAACTGTTTACCG 666  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 CGACGTTCCCGCTCGATTAAACCGGTAACCGCTGTTGCTAACAGAACTGTTTACCG 480  
QY 667 GCGTCACCGTTGTAC 681  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 GCGTCACCGTTGTAC 495

## RESULT 15

AAS90335  
ID AAS90335 standard; cDNA; 1207 BP.

XX AC

XX AAS90335;

XX 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #26139.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX P-PSDB; ABG26148.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX Claim 1; SEQ ID NO 26139; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (II) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1207 BP; 316 A; 265 C; 288 G; 338 T; 0 U; 0 Other;  
Query Match 69.1%; Score 472.8; DB 5; Length 1207;  
Best Local Similarity 99.2%; Pred. No. 3.6e-148;  
Matches 496; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
QY 187 ATGTACACACCCCGGTTAAACGCAACTGCGTGGCTGGCTATTTGTAAGAGATCCCTGCC 246  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
652 ATGTACACACCCCGGTTAAACGCAACTGCGTGGCTGGCTATTTGTAAGAGATCCCTGCC 711  
QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 306  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
712 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 771  
QY 307 CAGCAGGGAACAAACGGTTAAGCGTGACGATGATCCCGAAGGGTTTGTGAGCAGAAATGA 366  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
772 CAGCAGGGAACAAACGGTTAAGCGTGACGATGATCCCGAAGGGTTTGTGAGCAGAAATGA 831  
QY 367 ATCCCGGAAGAACTGCGCTGCTCCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTG 426  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
832 ATCCCGGAAGAACTGCGCTGCTCCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTG 891  
QY 427 T-TATGTGCCGATGGTGAACCGTGGCTTGGCC-GGTGCTACCGTCTGTTCTGTGCAACGT 484  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
892 TGTATGTGCCGATGGCGAACCCTGCGGGTCTGTAACCGTCTGTTCTGTGCAACGT 951  
QY 485 TAACGGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTATCTGT 544  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
952 TAACGGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTATCTGT 1011  
QY 545 TCACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGTGATCCCGGCTGTGGG 604  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1012 TCACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGTGATCCCGGCTGTGGG 1071  
QY 605 GCGACGTTCCCGCTCGATTAAAGCGGTAACCGCTGTTGCTAACAGAACTGTTTTTAC 664  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1072 GCGACGTTCCCGCTCGATTAAAGCGGTAACCGCTGTTGCTAACAGAACTGTTTTTAC 1131  
QY 665 CGGCGTCACCGTTGTACTAA 684  
Dd ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1132 CGGCGTCACCGTTGTACTAA 1151

Search completed: July 11, 2006, 22:41:55  
Job time : 590 secs



Result No.	Query #		DB	ID	Description
	Score	Match Length			
1	684	100.0	684	2	AR428736
2	684	100.0	684	2	AR455412
3	684	100.0	684	2	AX329368
4	684	100.0	684	2	AX382258
5	496	72.8	854	2	AR144788
6	498	72.8	1498	15	ECOUBIA
7	498	72.8	2000	2	E11274
8	498	72.8	2000	15	ECOUBICA
9	498	72.8	2000	15	ECOUBIPLS
10	498	72.8	2034	15	ECUBIAC
11	498	72.8	2348	15	ECUBI
12	498	72.8	6641	2	AR144789
13	498	72.8	110000	15	U00096 <sup>42</sup>
14	498	72.8	110000	15	AP009048 <sup>42</sup>
15	498	72.8	176195	15	ECOW89 <sup>42</sup>
16	496.4	72.6	1396	15	DQ087228
17	495	72.4	495	2	AR47932
18	495	72.4	495	2	AR428733
19	495	72.4	495	2	AR428736
20	495	72.4	495	2	AR455412
21	495	72.4	495	2	AX329368
22	495	72.4	495	2	AX382258
23	495	72.4	495	2	AR144788
24	495	72.4	495	15	M92668
25	495	72.4	495	15	Escherichia coli
26	495	72.4	495	15	Escherichia coli
27	495	72.4	495	15	Escherichia coli
28	495	72.4	495	15	Escherichia coli
29	495	72.4	495	15	Escherichia coli
30	495	72.4	495	15	Escherichia coli
31	495	72.4	495	15	Escherichia coli
32	495	72.4	495	15	Escherichia coli
33	495	72.4	495	15	Escherichia coli
34	495	72.4	495	15	Escherichia coli
35	495	72.4	495	15	Escherichia coli
36	495	72.4	495	15	Escherichia coli
37	495	72.4	495	15	Escherichia coli
38	495	72.4	495	15	Escherichia coli
39	495	72.4	495	15	Escherichia coli
40	495	72.4	495	15	Escherichia coli
41	495	72.4	495	15	Escherichia coli
42	495	72.4	495	15	Escherichia coli
43	495	72.4	495	15	Escherichia coli
44	495	72.4	495	15	Escherichia coli
45	495	72.4	495	15	Escherichia coli
46	495	72.4	495	15	Escherichia coli
47	495	72.4	495	15	Escherichia coli
48	495	72.4	495	15	Escherichia coli
49	495	72.4	495	15	Escherichia coli
50	495	72.4	495	15	Escherichia coli
51	495	72.4	495	15	Escherichia coli
52	495	72.4	495	15	Escherichia coli
53	495	72.4	495	15	Escherichia coli
54	495	72.4	495	15	Escherichia coli
55	495	72.4	495	15	Escherichia coli
56	495	72.4	495	15	Escherichia coli
57	495	72.4	495	15	Escherichia coli
58	495	72.4	495	15	Escherichia coli
59	495	72.4	495	15	Escherichia coli
60	495	72.4	495	15	Escherichia coli
61	495	72.4	495	15	Escherichia coli
62	495	72.4	495	15	Escherichia coli
63	495	72.4	495	15	Escherichia coli
64	495	72.4	495	15	Escherichia coli
65	495	72.4	495	15	Escherichia coli
66	495	72.4	495	15	Escherichia coli
67	495	72.4	495	15	Escherichia coli
68	495	72.4	495	15	Escherichia coli
69	495	72.4	495	15	Escherichia coli
70	495	72.4	495	15	Escherichia coli
71	495	72.4	495	15	Escherichia coli
72	495	72.4	495	15	Escherichia coli
73	495	72.4	495	15	Escherichia coli
74					

```
Db      181  TGGCATATGTACACCCCGCTTAACGCAACTGCGTGGCGCTATTGTAAGAGATC 240
Qy      241  CTTGCCCTGGATCCGAACTGCTCGACTGGCTGGCTGGAGGATTCATGACAAAACGT 300
Db      241  CTTGCCCTGGATCCGAACTGCTCGACTGGCTGGCTGGAGGATTCATGACAAAACGT 300
Qy      301  TTTGAAACAGCAGGAGGAAAACGGTAAGCGTACGATGATCCGGAAGGGTTTGTCCAGCAG 360
Db      301  TTTGAAACAGCAGGAGGAAAACGGTAAGCGTACGATGATCCGGAAGGGTTTGTCCAGCAG 360
Qy      361  AATGAAATCCCGGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAA 420
Db      361  AATGAAATCCCGGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAA 420
Qy      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTG 480
Db      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTG 480
Qy      481  ACGTTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
Db      481  ACGTTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
Qy      541  CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Db      541  CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Qy      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
Db      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy      661  TTACCGGCGTCAACCGTTGTAATAA 684
Db      661  TTACCGGCGTCAACCGTTGTAATAA 684
```

```
RESULT 2
LOCUS      AR455412                      684 bp      DNA      linear      PAT 20-FEB-2004
DEFINITION Sequence 7 from patent US 6683231.
ACCESSION  AR455412
VERSION    AR455412.1  GI:42689960
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 684)
AUTHORS   Meyer, K., Viitanen, P.V. and Van Dyk, D.E.
TITLE     High level production of p-hydroxybenzoic acid in green plants
JOURNAL   Patent: US 6683231-A 7 27-JAN-2004;
          E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES   Location/Qualifiers
            source
              1..684
              /organism="unknown"
              /mol_type="genomic DNA"
ORIGIN
  Query Match      100.0%; Score 684; DB 2; Length 684;
  Best Local Similarity 100.0%; Pred. No. 6.8e-202;
  Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Qy      1  ATGGCTTCCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
  Db      1  ATGGCTTCCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
  Qy      61  AGCATGGTTGCACTTCACTGGCTCTCAAAATTTTCAGCCACTTTCCCTGTTTACAAAGAG 120
  Db      61  AGCATGGTTGCACTTCACTGGCTCTCAAAATTTTCAGCCACTTTCCCTGTTTACAAAGAG 120
  Qy      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGATGCAAGTG 180
  Db      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGATGCAAGTG 180
  Qy      181  TGGCATATGTACACCCCGCTTAACGCAACTGCGTGGCGCTATTGTAAGAGATC 240
```

```
Db      181  TGGCATATGTACACCCCGCTTAACGCAACTGCGTGGCGCTATTGTAAGAGATC 240
Qy      241  CTTGCCCTGGATCCGAACTGCTCGACTGGCTGGCTGGAGGATTCATGACAAAACGT 300
Db      241  CTTGCCCTGGATCCGAACTGCTCGACTGGCTGGCTGGAGGATTCATGACAAAACGT 300
Qy      301  TTTGAAACAGCAGGAGGAAAACGGTAAGCGTACGATGATCCGGAAGGGTTTGTCCAGCAG 360
Db      301  TTTGAAACAGCAGGAGGAAAACGGTAAGCGTACGATGATCCGGAAGGGTTTGTCCAGCAG 360
Qy      361  AATGAAATCCCGGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAA 420
Db      361  AATGAAATCCCGGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAA 420
Qy      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTG 480
Db      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTG 480
Qy      481  ACGTTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
Db      481  ACGTTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAAAACGCCGTTAGGACGCTAT 540
Qy      541  CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Db      541  CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGATGCCGGGCTG 600
Qy      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
Db      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy      661  TTACCGGCGTCAACCGTTGTAATAA 684
Db      661  TTACCGGCGTCAACCGTTGTAATAA 684
```

RESULT 3

LOCUS AX329368 684 bp DNA linear PAT 08-JAN-2002

DEFINITION Sequence 7 from Patent WO0194607.

ACCESSION AX329368

VERSION AX329368.1 GI:18102367

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 other sequences; artificial sequences.

AUTHORS Meyer, K., van Dyk, D.E. and Viitanen, P.V.

TITLE High level production of p-hydroxybenzoic acid in green plants

JOURNAL Patent: WO 0194607-A 7 13-DEC-2001; E.I. Dupont De Nemours (US)

FEATURES Location/Qualifiers

source

1..684

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="synthetic CPL"

ORIGIN

Query Match 100.0%; Score 684; DB 2; Length 684;

Best Local Similarity 100.0%; Pred. No. 6.8e-202;

Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60

Db 1 ATGGCTTCCCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60

Qy 61 AGCATGGTTGCACTTCACTGGCTCTCAAAATTTTCAGCCACTTTCCCTGTTTACAAAGAG 120

Db 61 AGCATGGTTGCACTTCACTGGCTCTCAAAATTTTCAGCCACTTTCCCTGTTTACAAAGAG 120

Qy 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGATGCAAGTG 180

121	CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG	180
181	TGGCATATGTACACACCCCGCGTTAAACGCAACTGGCGTGCCTATTTGTTAAAGAGATC	240
181	TGGCATATGTACACACCCCGCGTTAAACGCAACTGGCGTGCCTATTTGTTAAAGAGATC	240
241	CTTGCCCTTGGANTCCGAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT	300
241	CTTGCCCTTGGANTCCGAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGT	300
301	TTTGAACAGCAGGAGGAAACCGTAAAGCTGACGATGATCCCGAAGGGTTTGTCCAGCAG	360
301	TTTGAACAGCAGGAGGAAACCGTAAAGCTGACGATGATCCCGAAGGGTTTGTCCAGCAG	360
361	AATGAAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA	420
361	AATGAAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA	420
421	ATTTTGTATTGTCCGATGTGTGAACCGTGGCTTGCCGGTGTGACCGTCTCTCTGTGCA	480
421	ATTTTGTATTGTCCGATGTGTGAACCGTGGCTTGCCGGTGTGACCGTCTCTCTGTGCA	480
481	ACGTTAAGCGGGCCGGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCCTAT	540
481	ACGTTAAGCGGGCCGGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCCTAT	540
541	CTGTTACATCATCGACATTAAACCGGACCTTTATTGAGATAGGCGCGTATGCCGGGCTG	600
541	CTGTTACATCATCGACATTAAACCGGACCTTTATTGAGATAGGCGCGTATGCCGGGCTG	600
601	TGGGGGCGACGTTCCCGCTCGGATTAAACGGTTAAACCGCTGTTGCTAACGAACCTGTTT	660
601	TGGGGGCGACGTTCCCGCTCGGATTAAACGGTTAAACCGCTGTTGCTAACGAACCTGTTT	660
661	TTACCGGGCGTACCGTTGTACTAA	684
661	TTACCGGGCGTACCGTTGTACTAA	684

RESULT_4			
AX382258			
LOCUS			
DEFINITION	684 bp	DNA	linear
ACCESSION	Sequence 16 from Patent WO0204653.		
VERSION	AX382258		
KEYWORDS	AX382258.1 GI:19577037		
SOURCE	.		
ORGANISM	synthetic construct		
	synthetic constructs		
	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Flint,D., Meyer,K. and Viitanen,P.V.		
TITLE	Sinapolyglucose:maltate sinapolytransferase form malate conjugates from benzic acid glucosides		
JOURNAL	Patent: WO 0204653-A 16 17-JAN-2002; E.I. DUPONT DE NEMOURS AND COMPANY (US)		
FEATURES	Location/Qualifiers		
source	1..684		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="open reading frame of the chloroplast-targeted CPL fusion protein"		

[illegible]



```

RESULT 7
E11274
LOCUS      ubiC and ubiA gene.                2000 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION
ACCESSION  E11274
VERSION    E11274.1 GI:22024916
KEYWORDS   JP 1996107789-A/1.
SOURCE     Escherichia coli
ORGANISM   Escherichia coli
Bacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE  1 (bases 1 to 2000)
AUTHORS    Matsuda,H., Kawamuki,M. and Nakagawa,T.
TITLE      PRODUCTION OF UBIQUINONE-10
JOURNAL    Patent: JP 1996107789-A 1 30-APR-1996;
ALPHA- SHOKUHIN KK
COMMENT    OS Escherichia coli
PN JP 1996107789-A/1
PD 30-APR-1996
PF 13-OCT-1994 JP 1994273071
PI MATSUDA HIDEYUKI, KAWAMUKI MAKOTO, NAKAGAWA TSUYOSHI PC
C12N15/09,A61K31/12,A61K31/12,C07H21/04,C12N1/21,C12P7/66, PC
(C12N1/21,
PC C12R1:01), (C12P7/66,C12R1:01);
CC Strandedness: Double;
CC topology: Linear;
FH Key      Location/Qualifiers
FH          1. .2000
FT source   /organism='Escherichia coli'
FT CDS      380..877
FT          /product='UbiC protein'
FT CDS      890..1762
FT          /product='UbiA protein'.
FEATURES    Location/Qualifiers
source      1. .2000
/mol_type="genomic DNA"
/db_xref="taxon:562"

ORIGIN
Query Match      72.8%; Score 498; DB 2; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTGTAAGAGATCCCTGCC 246
DB 380 ATGTCACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTGTAAGAGATCCCTGCC 439
QY 247 CTGGATCCGCAACTGCTCGA CTGGCTGTTCTGCTGGAGGATTCATGACAAACGTTTGAA 306
DB 440 CTGGATCCGCAACTGCTCGA CTGGCTGTTCTGCTGGAGGATTCATGACAAACGTTTGAA 499
QY 307 CACGAGGAAAAACGTTAAGCTGACATGATCCGGAAGGTTTTCGAGCAGAGATGAA 366
DB 500 CACGAGGAAAAACGTTAAGCTGACATGATCCGGAAGGTTTTCGAGCAGAGATGAA 559
QY 367 ATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTTG 426
DB 560 ATCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTTG 619
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTTCGCGGTGCTACCGTCTGTTCTGTGCAACGTTA 486
DB 620 TTATGTGCCGATGGTGAACCGTGGCTTTCGCGGTGCTACCGTCTGTTCTGTGCAACGTTA 679
QY 487 AGCGGCGGAGCTGGGTTACAAAATTTGGTTAAACCGCGTTAGACGCTATCTGTTTC 546
DB 680 AGCGGCGGAGCTGGGTTACAAAATTTGGTTAAACCGCGTTAGACGCTATCTGTTTC 739
QY 547 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 606
DB 740 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 799

```

```

QY 607 CGACGTTCCCGCTCGATTAAACCGGTAAACCGCTGCTGCTTAACAGAACTGTTTTACCG 666
DB 800 CGACGTTCCCGCTCGATTAAACCGGTAAACCGCTGCTGCTTAACAGAACTGTTTTACCG 859
QY 667 GCGTCACCGTTGTAATAA 684
DB 860 GCGTCACCGTTGTAATAA 877

RESULT 8
ECUBICA      2000 bp      DNA      linear      BCT 26-APR-1993
LOCUS      Escherichia coli chorismate lyase (ubiC), 4-hydroxybenzoate
DEFINITION  octaprenyl transferase (ubiA) genes, complete cds, and
ACCESSION  M93136
VERSION    M93136.1 GI:148099
KEYWORDS   4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plsB
SOURCE     Escherichia coli W3110
ORGANISM   Escherichia coli W3110
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE  1 (bases 1 to 2000)
AUTHORS    Lightner,V.A., Bell,R.M. and Modrich,P.
TITLE      The DNA sequences encoding plsB and dgk loci of Escherichia coli
JOURNAL    J. Biol. Chem. 258 (18), 10856-10861 (1983)
COMMENT    6309817
REFERENCE  2 (bases 1 to 2000)
AUTHORS    Nichols,B.P. and Green,J.M.
TITLE      Cloning and Sequencing of Escherichia coli ubiC and purification of
JOURNAL    chorismate lyase
COMMENT    Unpublished (1992)
FEATURES    Location/Qualifiers
source      1. .2000
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohare lambda-lf8"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorismate --> p-hydroxybenzoate +
pyruvate"
/experiment="experimental evidence, no additional details
recorded"
/citation=[2]
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA24711.1"
/db_xref="GI:148100"
/tranlation="MSHPALTOLRALRYCKEIPALDPOLLMLLEDSTMTRFEQOGK
TVSVNIRGFEQNEIPELLELLPKESRYWLEILLCADGEPLAGRTVVVPTLSC
PELAJQKGTPLGRYLFYTSITLTDIFEIGRDAGLWGRRLRLSGKPLLLLEFLP
ASPLY"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate
synthesis"
/experiment="experimental evidence, no additional details
recorded"
/citation=[2]

```

```

/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24712.1"
/db_xref="GI:148101"
/translacion="MEWSLTQNKLLAFHRLMRTDKP IGALLLLMPTLWALWVATPGVP
QLWILAVFVAGVWLMRAAGCVVNDYADRFKDFGHVKTANRPLPSGAVTEKEARALFVV
LVLSIFLLVLTNTWLTLLISAALAWVYPMKRYTHLPQVVLGAAGFWSIPMAFAA
VSESVPFLSCLMFLANILWAVADTOYAMVDRDDVKIGIKSTAILFGQYDKLILGIL
QIGVLALMIIIGELNGLGWYWSILVAGALFYQOKLIANRREACFKAFNNNYIVG
LVFLGLAWSYWHF"
/complement(1917..2000)
/gene="plsB"
/complement(1917..2000)
/gene="plsB"
/EC_number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details
recorded"
/citation={1}
/codon_start=1
/transl_table=11
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AAA24713.1"
/db_xref="GI:148102"
/translacion="MKVYQLLAELITSDVRLTIESATQEG"
```

## ORIGIN

```

Query Match          72.8%; Score 498; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTACACACCCCGTTACGCAACTCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 246
DB 380 ATGTACACACCCCGTTACGCAACTCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 439

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 306
DB 440 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 499

QY 307 CAGCAGGGAAAACCGTAAGCGTGAAGATGATCCGCGAGGTTTGTGAGCAGAAATGA 366
DB 500 CAGCAGGGAAAACCGTAAGCGTGAAGATGATCCGCGAGGTTTGTGAGCAGAAATGA 559

QY 367 ATCCCGGAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTAGCTGAAATTTTG 426
DB 560 ATCCCGGAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTAGCTGAAATTTTG 619

QY 427 TTATGTGCGGATGGTGAACCGTGGCTTCCCGGTGCTACCGTGGTTCCTGTGTCACGTTA 486
DB 620 TTATGTGCGGATGGTGAACCGTGGCTTCCCGGTGCTACCGTGGTTCCTGTGTCACGTTA 679

QY 487 AGCGGGCGGAGCTGGCTGACAAAATTTGGTAAACCGGTTAGGAGCTATCTGTTTC 546
DB 680 AGCGGGCGGAGCTGGCTGACAAAATTTGGTAAACCGGTTAGGAGCTATCTGTTTC 739

QY 547 ACATCATCGACATTAACCGGACTTTATTGAGATAGCGCGCTGATCGCGGCTGTGGGG 606
DB 740 ACATCATCGACATTAACCGGACTTTATTGAGATAGCGCGCTGATCGCGGCTGTGGGG 799

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTACC 666
DB 800 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTACC 859

QY 667 GCGTCACCGTTGTACTAA 684
DB 860 GCGTCACCGTTGTACTAA 877
```

## RESULT 9

```

EC0UBIPLS          2000 bp      DNA      linear      BCT 26-APR-1993
LOCUS              Escherichia coli 4-hydroxybenzoate octaprenyl transferase (ubiA)
```

```

ACCESSION
VERSION
KEYWORDS

SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

COMMENT

FEATURES
source

gene
CDS

gene
CDS
```

```

gene complete cds, chorismate lyase (ubiC) gene complete cds,
sn-glycerol-3-phosphate acyltransferase (plsB) gene, 3' end.
M93413.1 GI:148106
M93413.1 GI:148106
4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plsB
gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC
gene.
Escherichia coli W3110
Escherichia coli W3110
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 2000)
The DNA sequences encoding plsB and dgk loci of Escherichia coli
Lightner,V.A., Bell,R.M. and Modrich,P.
J. Biol. Chem. 258 (18), 10856-10861 (1983)
6309817
2 (bases 1 to 2000)
Nichols,B.P. and Green,J.M.
Cloning and sequencing of Escherichia coli ubiC and purification of
chorismate lyase
J. Bacteriol. 174 (16), 5309-5316 (1992)
1644758
Original source text: Escherichia coli (sub_strain W3110, strain
K-12) (library: Kohara lambda) DNA.
Location/Qualifiers
1..2000
/organism="Escherichia coli W3110"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohara lambda"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorismate --> p-hydroxybenzoate +
pyruvate"
experiment="experimental evidence, no additional details
recorded"
/note="N-terminal amino acid sequence of chorismate lyase
:NH2-SHPALTQRLRYCEIPAL-"
/citation={2}
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA24716.1"
/db_xref="GI:148107"
/translacion="MSHPALTQRLRYCKEIPALDQDLWLLLEDSMTKRFEQQKG
TVSVTMRGEFVEQNEIPELPLPKESRYWLRILLICADGEPWLAGRTVPVSTLSG
PELALOKLGTPLGRYLFSTSTLTRDFEIGRDAGLRGRRSRLSGKPLLLTELFLP
ASPLY"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate
synthesis"
experiment="experimental evidence, no additional details
recorded"
/citation={2}
/number=2
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24717.1"
/db_xref="GI:148108"
/translacion="MEWSLTQNKLLAFHRLMRTDKP ICAALLLLMPTLWALWVATPGVP
QLWILAVFVAGVWLMRAAGCVVNDYADRFKDFGHVKTANRPLPSGAVTEKEARALFVV
LVLSIFLLVLTNTWLTLLISAALAWVYPMKRYTHLPQVVLGAAGFWSIPMAFAA
VSESVPFLSCLMFLANILWAVADTOYAMVDRDDVKIGIKSTAILFGQYDKLILGIL
QIGVLALMIIIGELNGLGWYWSILVAGALFYQOKLIANRREACFKAFNNNYIVG
```

gene LVLFLGLAMSYWHP  
 complement (1917. .2000)  
 /gene="pleB"  
 CDS complement (1917. .2000)  
 /gene="pleB"  
 /EC\_number="2.3.1.15"  
 /function="lysophosphatidic acid synthesis"  
 /experiment="experimental evidence, no additional details  
 recorded"  
 /citation={1}  
 /codon\_start=1  
 /transl\_table=11  
 /product="sn-glycerol-3-phosphate acyltransferase"  
 /protein\_id="AA24718.1"  
 /db\_xref="GI:148109"  
 /translation="MKVYQLAELITSDVRLTIESATQEGG"

## ORIGIN

Query Match 72.8%; Score 498; DB 15; Length 2000;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-143; Indels 0; Gaps 0;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 246  
 Db 380 ATGTACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 439  
 QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 306  
 Db 440 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 499  
 QY 307 CAGCAGGGAACCGTAGCGTAGCATCCGCAAGGTTTCTCGAGCAGATGAA 366  
 Db 500 CAGCAGGGAACCGTAGCGTAGCATCCGCAAGGTTTCTCGAGCAGATGAA 559  
 QY 367 ATCCCCGAAGAACTGCGCTGCTGCCAAGAGTCTCGTTACTGGTTAGCTGAAATTTG 426  
 Db 560 ATCCCCGAAGAACTGCGCTGCTGCCAAGAGTCTCGTTACTGGTTAGCTGAAATTTG 619  
 QY 427 TTATGTCCGATGTTGAACCGTGGCTGTTGCGGTCGTTACCGTCTGTTGTCGCAAGTTA 486  
 Db 620 TTATGTCCGATGTTGAACCGTGGCTGTTGCGGTCGTTACCGTCTGTTGTCGCAAGTTA 679  
 QY 487 AGCGGCCCGAGCTGCGCTTACAAAATTTGGTAAACCGCTTAGCGCTATCTGTTTC 546  
 Db 680 AGCGGCCCGAGCTGCGCTTACAAAATTTGGTAAACCGCTTAGCGCTATCTGTTTC 739  
 QY 547 ACATCATGACATTACCGGCACTTTATTGATAGGCGCGTATCGCGGCTGTGGGG 606  
 Db 740 ACATCATGACATTACCGGCACTTTATTGATAGGCGCGTATCGCGGCTGTGGGG 799  
 QY 607 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTTACCG 666  
 Db 800 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTTACCG 859  
 QY 667 GCGTCACCGTTGACTAA 684  
 Db 860 GCGTCACCGTTGACTAA 877

## RESULT 10

ECUBIAC 2034 bp DNA linear BCT 05-OCT-1992  
 LOCUS E.coli ubiC and ubiA genes for chorismate lyase and  
 DEFINITION 4-hydroxybenzoate octaprenyltransferase.  
 X57434  
 VERSION X57434.1 GI:43233  
 KEYWORDS 4-hydroxybenzoate octaprenyl transferase; 4-hydroxybenzoate  
 synthetase; chorismate lyase.  
 SOURCE Escherichia coli  
 ORGANISM Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE 1

AUTHORS Nishimura, K., Nakahigashi, K. and Inokuchi, H.  
 TITLE Location of the ubiA gene on the physical map of Escherichia coli  
 J. Bacteriol. 174 (17), 5762 (1992)  
 PUBMED 1512213  
 REFERENCE 2 (bases 1 to 2034)  
 AUTHORS Nishimura, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1991) K. Nishimura, Faculty of Sciences, Kyoto  
 University, Sakyo-ku, Kyoto 606, Japan

FEATURES  
 Location/Qualifiers  
 1..2034  
 /organism="Escherichia coli"  
 /mol\_type="genomic DNA"  
 /strain="K12 W3110"  
 /db\_xref="taxon:562"  
 /map="91.7 min"  
 /clone="634"  
 /clone\_lib="Kohara library"  
 137..142  
 -35 signal  
 139..164  
 -10 signal  
 291..899  
 /gene="ubiC"  
 291..899  
 /gene="ubiC"  
 /codon\_start=1  
 /transl\_table=11  
 /product="4-hydroxybenzoate synthetase"  
 /protein\_id="CAA40681.1"  
 /db\_xref="GI:43234"  
 /db\_xref="GOA:P26602"  
 /db\_xref="UniProtKB/Swiss-Prot:P26602"  
 /translation="MRLRPFCCVLDHLICFTSPVNTFLRYNAFTLNCBFGMSHPALT  
 QURALRYKEIPALDPQLDMLLEDMSMTREFQOQKTVSTMIRGEFVEQNSIPEL  
 PLPLKESRYLWREILLCADGEPLAGRTVVPVSTLSPPELALQKLKGLTPLGRYLFST  
 TLTRDFEIGRDAGLWRRSRRLSGKPLLLTELFLPASPLY"

gene  
 CDS

912..1784  
 /gene="ubiA"  
 912..1784  
 /gene="ubiA"  
 /codon\_start=1  
 /transl\_table=11  
 /product="4-hydroxybenzoate octaprenyltransferase"  
 /protein\_id="CAA40682.1"  
 /db\_xref="GI:43235"  
 /db\_xref="GOA:P26601"  
 /db\_xref="UniProtKB/Swiss-Prot:P26601"  
 /translation="MEWSLTQKLLAFHRLMRDTPIGALLLLWPTLWALWVATPGVP  
 OLWILAVFVAGVWLMRAAGCVNDVADKFDGKVRKTNRLPLPSGAVTEKEARALFV  
 LVLISFLVLTLNTWTILLIAALAMVYPEWKRYTHLPQVVLGAAPGWSIPMAFAA  
 VSESVPLSCHWLMFLANIWNAYDQYANVDRDDVKIGIKSTAILFGQYDKLIIIGIL  
 QIGVLAALWIIIGELNGWGYYSILVAGALFVYQOKLIANREREAQFKAFMNNVTYV  
 LVLFLGLAMSYWHP"

## ORIGIN

Query Match 72.8%; Score 498; DB 15; Length 2034;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-143; Indels 0; Gaps 0;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 246  
 Db 402 ATGTACACCCCGCTTAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 461  
 QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 306  
 Db 462 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 521  
 QY 307 CAGCAGGGAACCGTAGCGTAGCATCCGCAAGGTTTCTCGAGCAGATGAA 366  
 Db 522 CAGCAGGGAACCGTAGCGTAGCATCCGCAAGGTTTCTCGAGCAGATGAA 581  
 QY 367 ATCCCCGAAGAACTGCGCTGCTGCCAAGAGTCTCGTTACTGGTTAGCTGAAATTTG 426  
 Db 582 ATCCCCGAAGAACTGCGCTGCTGCCAAGAGTCTCGTTACTGGTTAGCTGAAATTTG 641









Db 50949 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTTTACCG 51008  
Qy 667 GCGTCACCGTTGTACTAA 684  
Db 51009 GCGTCACCGTTGTACTAA 51026

RESULT 14  
AP009048\_42  
WPCOMMENT

Sequence split into 47 fragments LOCUS AP009048 Accession AP009048

Fragment Name	Begin	End
AP009048_00	1	110000
AP009048_01	100001	210000
AP009048_02	200001	310000
AP009048_03	300001	410000
AP009048_04	400001	510000
AP009048_05	500001	610000
AP009048_06	600001	710000
AP009048_07	700001	810000
AP009048_08	800001	910000
AP009048_09	900001	1010000
AP009048_10	1000001	1110000
AP009048_11	1100001	1210000
AP009048_12	1200001	1310000
AP009048_13	1300001	1410000
AP009048_14	1400001	1510000
AP009048_15	1500001	1610000
AP009048_16	1600001	1710000
AP009048_17	1700001	1810000
AP009048_18	1800001	1910000
AP009048_19	1900001	2010000
AP009048_20	2000001	2110000
AP009048_21	2100001	2210000
AP009048_22	2200001	2310000
AP009048_23	2300001	2410000
AP009048_24	2400001	2510000
AP009048_25	2500001	2610000
AP009048_26	2600001	2710000
AP009048_27	2700001	2810000
AP009048_28	2800001	2910000
AP009048_29	2900001	3010000
AP009048_30	3000001	3110000
AP009048_31	3100001	3210000
AP009048_32	3200001	3310000
AP009048_33	3300001	3410000
AP009048_34	3400001	3510000
AP009048_35	3500001	3610000
AP009048_36	3600001	3710000
AP009048_37	3700001	3810000
AP009048_38	3800001	3910000
AP009048_39	3900001	4010000
AP009048_40	4000001	4110000
AP009048_41	4100001	4210000
AP009048_42	4200001	4310000
AP009048_43	4300001	4410000
AP009048_44	4400001	4510000
AP009048_45	4500001	4610000
AP009048_46	4600001	4646332

Continuation (43 of 47) of AP009048 from base 4200001 (AP009048 Escherichia coli W3110

Query Match 72.8%; Score 499; DB 15; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 3e-143;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 187 ATGTACACCCCGCTTAACGCACTGCGTGGCGCTGCGCTATGTAAAGAGATCCCTGCC 246  
Db 56096 ATGTACACCCCGCGCTTAACGCACTGCGTGGCGCTATGTAAAGAGATCCCTGCC 56155  
Qy 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATCACAAAAAGCTTTTGA 306  
Db 56156 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATCACAAAAAGCTTTTGA 56215

Qy 307 CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGGAAGGGTTTGTGCGAGCAGATGAA 366  
Db 56216 CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGGAAGGGTTTGTGCGAGCAGATGAA 56275  
Qy 367 ATCCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGTTACGTGAAATTTTG 426  
Db 56276 ATCCCCGAAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGTTACGTGAAATTTTG 56335  
Qy 427 TTATGTGCCGATGGTGAACCGTGGCTTCCCGTGTACCGTCTGTTCCCTGTGTCAACGTTA 486  
Db 56336 TTATGTGCCGATGGTGAACCGTGGCTTCCCGTGTACCGTCTGTTCCCTGTGTCAACGTTA 56395  
Qy 487 AGCGGGCCGAGCTGGCGTTACAAAAATTTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 546  
Db 56396 AGCGGGCCGAGCTGGCGTTACAAAAATTTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 56455  
Qy 547 ACATCATCGACATTAAACCGGAGCTTTATTAGATAGCCGCTGATGCCGGCTGTGGGG 606  
Db 56456 ACATCATCGACATTAAACCGGAGCTTTATTAGATAGCCGCTGATGCCGGCTGTGGGG 56515  
Qy 607 CGAGTTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAAACAGAACTGTTTTACCG 666  
Db 56516 CGAGTTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAAACAGAACTGTTTTACCG 56575  
Qy 667 GCGTCACCGTTGTACTAA 684  
Db 56576 GCGTCACCGTTGTACTAA 56593

RESULT 15

ECOW89 176195 bp DNA linear BCT 17-DEC-1993  
LOCUS E. coli chromosomal region from 89.2 to 92.8 minutes.  
DEFINITION U00006  
ACCESSION U00006.1 GI:409785  
KEYWORDS Escherichia coli  
SOURCE Escherichia coli  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE 1 (bases 1 to 176195)  
Blattner,F.R., Burland,V., Plunkett,G. III, Sofia,H.J. and  
Daniels,D.L.  
Analysis of the Escherichia coli genome. IV. DNA sequence of the  
region from 89.2 to 92.8 minutes  
Nucleic Acids Res. 21 (23), 5408-5417 (1993)  
PUBMED 8265357  
REFERENCE 2 (bases 1 to 176195)  
Blattner,F.R.  
Direct Submision  
Submitted (03-SEP-1993) 608-263-7459  
On Oct 29, 1993 this sequence version replaced gi:396288.  
This sequence was determined as part of the E. coli Genome Project  
(Frederick R. Blattner, director) at the University of  
Wisconsin-Madison. Supported by award HG00301 from the NIH Human  
Genome Project. The entire sequence was independently determined  
from E. coli MGI655; overlaps and conflicts with other sequence  
determinations are annotated. The start of this entry overlaps the  
end of the entry ECOW87 (U19201) by 93 bp. NOTE: An update was  
submitted on 25-OCT-1993, reflecting a correction to the heme  
sequence.

FEATURES  
source

Location/Qualifiers  
1..176195  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"  
/strain="K-12"  
/sub\_strain="MG1655"  
/db\_xref="taxon:562"

/note="This sequence comprises the following lambda  
clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203,  
EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262,  
EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27;  
M13mp19 or Janus vectors were used for subcloning"



```
CDS
7771..8850
/note="similar to phosphotransferase system enzyme II"
/codon_start=1
/translation_table=11
/label=ORF_0359
/protein_id="AAC43055.1"
/db_xref="GI:396296"
/translation="MNEIVQILKNTROHMTGVSHMIPFVWSGGILLAVSVMLYKGA
VPDAVDADNLKLPDIFGVAGLTLMPFLAAYIGVSIABERSALAPCAIGRWVNSFGAG
FFGALIAGLIGIVVHDKPIPVHKLRSVMPFIPIPIVGLITAGIMMGLGEPVGA
LTSUTQMLQMGQGSIVMLAIVIMGLAFDMGPEVKNVAYAFMLICVAQGYVTVAI
AAVGICIPPLGLMLATLIGRKNFSABERETKAALVMGCVGVTEGAIPFAAADPLRVI

Query Match      72.8%; Score 498; DB 15; Length 176195;
Best Local Similarity 100.0%; Pred. No. 3.4e-143; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 246
Db 117748 ATGTCACACCCCGCGTTAACGCAACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 117807

QY 247 CTGGATCCGCAACTGCTGACTGGCTGTTCTGGAGGATTCATGACAAAACGTTTGGAA 306
Db 117808 CTGGATCCGCAACTGCTGACTGGCTGTTCTGGAGGATTCATGACAAAACGTTTGGAA 117867

QY 307 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCCGAAGGGTTTGTGAGCAGAAATGAA 366
Db 117868 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCCGAAGGGTTTGTGAGCAGAAATGAA 117927

QY 367 ATCCCGGAAGAACTCCCGTGTGTCGCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTTG 426
Db 117928 ATCCCGGAAGAACTCCCGTGTGTCGCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTTG 117987

QY 427 TTATGTGCGGATGGTGAACCGTGGCTTGCCGGTCTGACCGTGGTTCCTGTGTCAACGTTA 486
Db 117988 TTATGTGCGGATGGTGAACCGTGGCTTGCCGGTCTGACCGTGGTTCCTGTGTCAACGTTA 118047

QY 487 AGCGGGCCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGCGTTAGGACGCTATCTGTTTC 546
Db 118048 AGCGGGCCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCGCGTTAGGACGCTATCTGTTTC 118107

QY 547 ACATCATCGACATTAAACCGGACTTTTATTGAGATAGCCGGTGTGATGCCGGGCTGTGGGGG 606
Db 118108 ACATCATCGACATTAAACCGGACTTTTATTGAGATAGCCGGTGTGATGCCGGGCTGTGGGGG 118167

QY 607 CGACGTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 666
Db 118168 CGACGTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 118227

QY 667 GCGTCACCGTTGTACTAA 684
Db 118228 GCGTCACCGTTGTACTAA 118245
```

Search completed: July 11, 2006, 20:23:15  
Job time : 4397 secs

Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 100.0%  
DB: 2  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-718-311-15 (1-62) x AX382258 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 1 ATGGCTTCCTCTGTCATTTCTTACAGCAGCTGTTGCCACACGACCAATGTTACACAGCT 60  
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 61 AGCATGGTTCACCTTTCTCAATGCTAGCAATGTTGGAAGAGTTCCTGTCATGAGGTG 120  
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerSerAlaThrArgValSerCysMetGlnVal 60  
Db 121 CAAACCTTGACATCTTCAATGCTAGCAATGTTGGAAGAGTTCCTGTCATGAGGTG 180

Qy 61 TrpHis 62  
Db 181 TGGCAT 186

RESULT 5  
TOMRBCSB 599 bp mRNA linear PLN 27-APR-1993  
LOCUS Tomato (L. esculentum) ribulose-1,5-bisphosphate  
DEFINITION carboxylase (RBCS) small subunit mRNA, clone Rbcs-2A.  
ACCESSION M13543  
VERSION M13543.1 GI:170497  
KEYWORDS ribulose 1,5-bisphosphate carboxylase/oxygenase.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
ORGANISM Lycopersicon esculentum

REFERENCE  
AUTHORS Pichersky, E., Tanksley, R., Tanksley, S.D. and Cashmore, A.R.  
TITLE Evidence for selection as a mechanism in the concerted evolution of  
Lycopersicon esculentum (tomato) genes encoding the small subunit  
of ribulose-1,5-bisphosphate carboxylase/oxygenase  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (11), 3880-3884 (1986)  
PUBMED 3012537  
COMMENT Original source text: Tomato (L. esculentum) cDNA to mRNA, clone  
Rbcs-2A.  
Clean copy sequence for [1] kindly provided by E. Pichersky,  
31-OCT-1986.

FEATURES  
source  
1..599  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4081"  
CDS  
15..557  
/note="ribulose-1,5-bisphosphate carboxylase/ oxygenase  
small subunit (EC 4.1.1.39)"  
/codon\_start=1  
/protein\_id="AAA34189.1"  
/db\_xref="GI:170498"  
/translation="MASSVISAATVATSNVTQASVAPFTGLKSSATFPVTKQNLD  
ITSASNGRVSQVMPPIINMKYETLSYLPDLSEQLLSVEYLLKNGWVPCLEFE  
TEHGFVYRENNKSPGYDGRYTWKLPFGCTDTQVLAIEVQEAQKAYPOAWVRIG  
FDNRVQVCISFIAYKPEGY"

ORIGIN  
Chromosome 3, Rbcs-2 locus.  
Alignment Scores:  
Pred. No.: 3,99e-28 Length: 599  
Score: 297.00 Matches: 61  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 97.4% Indels: 0  
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSB (1-599)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20  
Db 15 ATGGCTTCCTCTGTCATTTCTTACAGCAGCTGTTGCCACACGACCAATGTTACACAGCT 74  
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40  
Db 75 AGCATGGTTCACCTTTCTCAATGCTAGCAATGTTGGAAGAGTTCCTGTCATGAGGTG 134  
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerSerAlaThrArgValSerCysMetGlnVal 60  
Db 135 CAAACCTTGACATCTTCAATGCTAGCAATGTTGGAAGAGTTCCTGTCATGAGGTG 194  
Qy 61 Trp 61  
Db 195 TGG 197

RESULT 6  
TOMRBCSE 742 bp mRNA linear PLN 27-APR-1993  
LOCUS Tomato RuBP carboxylase small subunit LESS5 mRNA.  
DEFINITION M15236  
ACCESSION M15236.1 GI:170503  
KEYWORDS carboxylase; ribulose bisphosphate carboxylase.  
SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
ORGANISM Lycopersicon esculentum

REFERENCE  
AUTHORS McKnight, T.D., Alexander, D.C., Babcock, M.S. and Simpson, R.B.  
TITLE Nucleotide sequence and molecular evolution of two tomato genes  
encoding the small subunit of ribulose-1,5-bisphosphate carboxylase  
Gene 48 (1), 23-32 (1986)  
PUBMED 3557127  
COMMENT Original source text: Tomato (VF36), cDNA to mRNA, clone LESS5.  
Draft entry and computer-readable copy of sequence in [1] kindly  
provided by T.D. McKnight (15-APR-1987).

FEATURES  
source  
1..742  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4081"  
mRNA  
CDS  
28..570  
/product="RuBP mRNA"  
/note="ribulose-1,5-bisphosphate carboxylase, small  
subunit precursor"  
/codon\_start=1  
/protein\_id="AAA34192.1"  
/db\_xref="GI:170504"  
/translation="MASSVISAATVATSNVTQASVAPFTGLKSSATFPVTKQNLD  
ITSASNGRVSQVMPPIINMKYETLSYLPDLSEQLLSVEYLLKNGWVPCLEFE  
TEHGFVYRENNKSPGYDGRYTWKLPFGCTDTQVLAIEVQEAQKAYPOAWVRIG  
FDNRVQVCISFIAYKPEGY"

sig\_peptide  
mat\_peptide  
ORIGIN 614 bp upstream of DraI site.  
Alignment Scores:  
Pred. No.: 5,27e-28 Length: 742  
Score: 297.00 Matches: 61  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 97.4% Indels: 0  
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSE (1-742)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

**THIS PAGE BLANK (USPTO)**